OPTO-ELECTRONIC SYSTEMS RESEARCH

TECHNICAL REPORT

AD-A203 778

DIFFRACTION THEORY FOR

POLYGONAL APERTURES

R. Edward English, Jr.

DTIC
ELECTE
JAN 1 8 1989

July 1988

The Institute of Optics
University of Rochester

Prepared for:

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
Unclassified	

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
	ARO 24626.57-PH-UIR

6a. NAME OF PERFORMING ORGANIZATION	6b OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
University of Rochester		U. S. Army Research Office

6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
The Institute of Optics Rochester, New York 14627	P. O. Box 12211 Research Triangle Park, NC 27709-2211

8a. NAME OF FUNDING / SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
U. S. Army Research Office		*DAAL03-86-K-0173*

8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS			
P. O. Box 12211 Research Triangle Park, NC 27709-2211	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)

Diffraction Theory for Polygonal Apertures

12 PERSONAL AUTHOR(S)   R. Edward English, Jr.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical	FROM _____ TO _____	July 1988	186

16. SUPPLEMENTARY NOTATION   The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation

17.	COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Electromagnetic diffraction; Fourier analysis; Gaussian beam

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

We explain and describe diffraction from polygonal apertures over a wide range of sizes and observation distances. In the first case considered, a small square aperture ($2a \times 2a$, $ka \ll 1$, where $k = 2\pi/\lambda$ is the wavenumber) in a perfectly conducting plane screen of vanishing thickness diffracts a normally incident, linear polarized, monochromatic plane wave. Within the vector framework of Maxwell's equations, we hypothesize a solution for the dominant component of the electric field. Subsequently, by means of an integro-differential equation formulation of the diffraction problem applied to small apertures, we substantiate the solution. The solution represents the first three terms in a more general expansion for the aperture field. Physical intuition and the solutions for

. . . continued on next page

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Nicholas George	716-275-2417	

DD FORM 1473, 84 MAR   83 APR edition may be used until exhausted.   SECURITY CLASSIFICATION OF THIS PAGE
All other editions are obsolete.

19.   ABSTRACT continued from previous page . . .

circular apertures and slits motivate us to propose this expansion. Numerical
calculations validate the solution over most of the aperture except in the close
vicinity of the corners of the aperture. This limited expansion does not achieve
an accurate description of the field near the corners.

In the remainder of the investigation we treat diffraction within the realm of
Fourier optics. We develop a Gaussian beam expansion and is it to describe
diffraction from a plane-screen form this corner. For Gaussian illumination of
the corner, we consider several opening angles and explain computer-generated
plots of the diffracted intensities. In addition we compute the optical transform
of a uniformly illuminated triangle aperture and a sector of a circular aperture.
Both of these aperture are plane-screen corners, but they have different
bounding edges.

The triangle aperture transform solution constitutes a basis for describing
diffraction by a polygonal aperture. A suitable combination of rotated,
elemental triangular apertures can represent the polygon. Hence, the
diffraction pattern for the polygon comprises the diffraction patterns of the
elemental triangles. A similar decomposition procedure is the key to writing a
simple closed-form solution for diffraction by nested polygonal apertures. We
present results for the particular case of a regular pentagon; here, the elemental
building blocks are isosceles triangles with a grating-like structure. The
diffraction patterns of these nested apertures contain interesting, low intensity
features: nested polygons are traced out in the diffraction pattern. Numerical
calculations and careful measurements confirm them.

Accesion For		
NTIS CRA&I		☑
DTIC TAB		☐
Unannounced		☐
Justification		
By		
Distribution		
Availability Codes		
Dist	Avail and / or Special	
A-1		

DIFFRACTION THEORY FOR POLYGONAL APERTURES

by

R. Edward English Jr.

Submitted in Partial Fulfillment

of the

Requirements for the Degree

DOCTOR OF PHILOSOPHY

Supervised by Dr. Nicholas George

The Institute of Optics

University of Rochester

Rochester, New York

1988

VITAE

The author ███████ Ronald Edward English Jr. ████████████████
████████████ He studied at Purdue University in West Lafayette,
Indiana from 1978–1982 and received a B. S. in honors physics and
mathematics in May 1982. He graduated with highest distinction and was
a member of Phi Beta Kappa. His graduate work towards a Ph. D. in optics
began in the fall of 1982 at the University of Rochester. As a student in
The Institute of Optics he was an Institute fellow for his first year. He was
awarded an Amoco Foundation doctoral fellowship in the fall of 1984; that
support continued until graduation. He was also supported by the ARO-
URI program (Army Research Office's University Research Initiative).
Since the spring of 1983 he has been supervised by Professor Nicholas
George in the Opto-Electronic Systems Group.

ACKNOWLEDGEMENTS

DIFFRACTION THEORY FOR POLYGONAL APERTURES

ABSTRACT

We explain and describe diffraction from polygonal apertures over a wide range of sizes and observation distances. In the first case considered, a small square aperture ($2a \times 2a$, $ka \lesssim 1$, where $k = 2\pi/\lambda$ is the wavenumber) in a perfectly conducting plane screen of vanishing thickness diffracts a normally incident, linear polarized, monochromatic plane wave. Within the vector framework of Maxwell's equations, we hypothesize a solution for the dominant component of the electric field. Subsequently, by means of an integro-differential equation formulation of the diffraction problem applied to small apertures, we substantiate the solution. The solution represents the first three terms in a more general expansion for the aperture field. Physical intuition and the solutions for circular apertures and slits motivate us to propose this expansion. Numerical calculations validate the solution over most of the aperture except in the close vicinity of the corners of the aperture. This limited expansion does not achieve an accurate description of the field near the corners.

In the remainder of the investigation we treat diffraction within the realm of Fourier optics. We develop a Gaussian beam expansion and use it to describe diffraction from a plane-screen corner of arbitrary angle. Two intersecting, coplanar half-plane screens form this corner. For Gaussian illumination of the corner, we consider several opening angles and explain computer-generated plots of the diffracted intensities. In addition we compute the optical transform of a uniformly illuminated triangular

aperture and a sector of a circular aperture. Both of these aperture are plane-screen corners, but they have different bounding edges.

The triangular aperture transform solution constitutes a basis for describing diffraction by a polygonal aperture. A suitable combination of rotated, elemental triangular apertures can represent the polygon. Hence, the diffraction pattern for the polygon comprises the diffraction patterns of these elemental triangles. A similar decomposition procedure is the key to writing a simple closed-form solution for diffraction by nested polygonal apertures. We present results for the particular case of a regular pentagon; here, the elemental building blocks are isosceles triangles with a grating-like structure. The diffraction patterns of these nested apertures contain interesting, low-intensity features: nested polygons are traced out in the diffraction pattern. Numerical calculations and careful measurements confirm them.

TABLE OF CONTENTS

APPENDIX

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

In this thesis we present diffraction solutions for polygonal apertures
in two size domains (small and large compared to the wavelength of light
λ) and over a wide range of observation distances. For the purpose of
mathematical tractability, we assume the apertures to be holes in a
perfectly conducting, plane screen of vanishing thickness. In the case of
large apertures, we assume that the true electric field in the aperture is
approximated by the incident electric field. This assumption combined
with an approximation about the phase of the Green's function (Fresnel
approximation) places that analysis within the scope of Fourier optics. On
the other hand small apertures cannot be treated this way; the true
aperture electric field deviates significantly from the incident electric
field. A more complete analysis within the vector framework of Maxwell's
equations is necessary. We describe such an analysis.

1.1 Statement of the Problem

The investigation separates logically into two divisions: diffraction
theory for small apertures (Chapter 2) and diffraction theory for large
apertures (Chapters 3 and 4). The analysis in Chapter 2 relies on only the
idealizations of a perfectly conducting screen of vanishing thickness and an
aperture whose characteristic dimensions are much smaller than the
wavelength of light λ. The problem can be stated as:

> Determine the diffracted fields produced when a small,
> polygonal aperture – square aperture ($2a \times 2a$, $ka \ll 1$, where k

1

$= 2n/\lambda$ is the wavenumber) – in a thin, perfectly conducting plane screen obstructs a normally incident, linear polarized plane wave.

We formulate this unsolved diffraction problem in terms of a magnetic vector potential. This vector potential, which depends on the tangential components of the true aperture electric field, determines the electric and magnetic fields everywhere. Symmetry properties of the electromagnetic field and continuity requirements lead to a set of three coupled integro-differential equations for the aperture electric field. An edge condition on the behavior of the electric field along the rim of the aperture provides a physical framework for determining a solution. Finally, we simplify these equations by restricting attention to apertures much smaller than the wavelength of light.

It is then appropriate to review briefly the solution to a similar fundamental diffraction problem – a small circular aperture of radius a – within this framework. This provides a foundation for hypothesizing an approximate solution for the small square aperture problem. First, we present the solution as a general expansion for the aperture electric field. Retention of the first three terms in the series for the dominant component of the electric field constitutes the hypothesized solution. Then, we use analytic and numeric means to determine the three unknown coefficients.

The integro-differential equation framework provides a means for evaluating the validity of the hypothesized aperture field solution. The solution is further compared to published numeric results for the effective dipole moment of a small square aperture and to the precise aperture electric field solution for a small circular aperture. Finally, a discussion of

the far field diffraction pattern from small apertures makes clear the applicability and usefulness of the results.

The formula developed in Chapter 2 is a concise and informative mathematical description for the dominant component of the aperture field of this unsolved electromagnetic diffraction problem. The formula describing this aperture field has a simple structure, namely, a polynomial times the square root of a rational function. Thus, we can easily evaluate the field at any point within the aperture. Further, the nature of the electric field behavior near sharp edges is displayed explicitly and recognized readily in the formula. These two properties (mathematical simplicity and physical origin) increase the usefulness of the hypothesized solution for describing the diffracted fields in the far field and closer than the far field.

On the other hand, we recognize that there is a basic difficulty in posing this problem in a precise fashion since the behavior of the electric field in the near vicinity of the corners of the aperture is unknown. [Many investigators have considered the corner behavior (Jones, 1952; Braunbek, 1956), and this behavior is still a subject of current interest (De Smedt and Van Bladel, 1987).] Hence, we can only hypothesize that the solution in Chapter 2 approximates the true aperture fields in a small square aperture. Knowledge of this corner behavior and a statement of a corner condition for the exact solution is required to analyze fully any proposed solution. Nevertheless, on the basis of the integro-differential equation formulation and subsequent analysis, we conclude that the behavior of the electric field in the near vicinity of the corners of the aperture is not

represented well. Possibly, inclusion of higher order terms could enable one to describe more accurately the electric field near the corners.

The problem statement for the large aperture diffraction theory reads:

> Determine the optical transform of a set of nested regular polygons with the special intent of explaining the low-intensity features that have not been described or explained previously.

We achieve the solution in several steps. First, we study the diffraction pattern produced when an incident light wave encounters a pair of coplanar, intersecting half-plane screens, i.e., a corner of arbitrary angle, in Chapter 3. We describe a Gaussian beam expansion approach that is suitable for this diffraction problem. The expansion is based upon Hermite-Gaussian and Laguerre-Gaussian functions, which constitute a complete and orthogonal set of eigenfunctions for the paraxial scalar Helmholtz equation. Since the Fresnel diffraction integral is a solution to this equation, and each basis function retains its functional form under Fresnel propagation, expansion of an aperture distribution in terms of these eigenfunctions constitutes a solution of the diffraction problem as well. We carry out the explicit calculation of the expansion coefficients, and we explain computer calculations for plane-screen corners of various opening angles under Gaussian illumination. The general features of the calculation are confirmed by comparison to experimental photographs. In addition careful analysis and examination of the Gaussian beam expansion calculations for particular corners reveal previously unobserved diffraction phenomena.

We perform a direct calculation of the Fourier transform of a uniformly illuminated but finite or bounded plane-screen corner also. We consider two cases of interest: a slit boundary and a circular boundary. The first case, which is also referred to as a triangular aperture or a triangular section of a circular aperture, can be evaluated in a straightforward manner; the resulting formula is graphically displayed as a computer-generated diffraction pattern. We use this calculation as the framework for studying polygonal apertures in Chapter 4. The circular boundary problem, which can be considered diffraction from a sector of a circular aperture, is solved by expanding the integrand in an infinite series of Bessel functions. The solution is suitable for computer calculation and completes the study of plane-screen corners.

In Chapter 4 we address the large polygonal aperture diffraction problem directly. In a straightforward manner, we review how a polygon diffraction pattern is composed of triangle diffraction patterns; these triangles are the bounded corners studied in Chapter 3. Recognizing that a similar decomposition procedure can be applied to the nested polygon diffraction problem is the *coup de maître* that enables us to write a simple closed-form solution for the diffraction pattern. A triangular aperture with a grating-like structure is the elemental building block. Each elemental triangle generates intensity spikes that are replicated in a regular array by the grating-like structure. The arrangement of these spikes produces nested polygons in the diffraction pattern. These features are demonstrated clearly by computer calculations and photographs for nested pentagons. Experimental measurement of these features, which

are many orders of magnitude lower in intensity than the central peak intensity, verify the calculations.

1.2 Historical Overview

Extensive literature reviews of research on the general problem of electromagnetic diffraction by apertures are contained in papers by Bouwkamp (1954), Eggimann (1961), and Butler, Rahmat-Samii, and Mittra (1978). The following cited publications deal with diffraction from apertures that are smaller than or equal to a wavelength of light in size.

Lord Rayleigh investigated the interaction of waves with plane screens perforated by some kind of aperture. It is of particular relevance that he considered one or both dimensions of the aperture to be infinitely small in comparison with the wavelength of light λ. Specifically, Rayleigh (1897a) solved integral equations applicable to acoustic diffraction by small circular holes and disks and narrow slits and blades. For example, he showed that the amplitude of the wave diffracted by a small hole in a fixed screen is approximately equal to the electrostatic capacity of a metal disk of the same size and shape as the aperture. Subsequently, he considered the specific electromagnetic problem of the normal incidence of a plane wave upon an infinitely thin screen with perfect electrical conductivity, perforated by a small circular aperture (1897b). His solution at a distance great in comparison to the wavelength of light is in the form of the far field radiation pattern of a magnetic dipole. Rayleigh expressed concern in applying scalar results to these electromagnetic situations because of the discontinuities which occur at the edge of the aperture.

Bethe (1944) studied small hole electromagnetic diffraction problems and the effect of a small hole in a cavity upon the oscillation of that cavity. His method was based on the use of fictitious magnetic charges and currents in the hole to satisfy the boundary conditions. Specifically, he used the solution of an electrostatics potential problem to determine the charge density in a small circular hole; the equation of continuity then yielded the current density. Further, he showed that the diffracted field can be considered as caused by a magnetic moment in the plane of the hole and an electric moment perpendicular to it.

An integral equation method of solving plane diffraction problems was presented by Copson (1946). He expressed the components of the field as integrals over either the screen or the aperture; the components were coupled through Maxwell's equations. Copson applied his method to rederive Bethe's expressions for the field in a small circular aperture and also obtained a solution for the case of a disk.

Although both of the two preceding works contributed valuable insight and understanding, they are open to criticism and require qualification. Bouwkamp perceived and explained the inconsistencies. In his critique of Bethe's paper (Bouwkamp, 1950a), Bouwkamp pointed out that in Bethe's solution the electric field is discontinuous in the hole because the expression for the magnetic current density is in error. The error does not, however, invalidate the usefulness of Bethe's effective dipole moment concept in the distant field, although the field in and near the hole is wrong. Bouwkamp (1950b) presented the correct fields up to the sixth order.

In a review of Copson's paper (1946), Bouwkamp (1947) showed that Copson's proof of his theorem 4 was incomplete and required an extra condition regarding the behavior of the electric field on the rim of the hole. This inadequacy was described more fully later (Bouwkamp, 1953). In a subsequent paper Copson (1950) objected to this point, but he later withdrew the objection (1951). In any event, the simpler system of differential-integral equations introduced in Copson's later paper (1950) is sound when Bouwkamp's additional constraint is employed.

Meixner and Andrejewski (1950) derived a solution for the diffraction of electromagnetic waves by perfectly conducting disks and circular apertures in conducting plane screens. Using a Hertz potential representation, they expressed the field in a suitable eigenfunction expansion. Since the boundary conditions are specified easily in an oblate spheroidal coordinate system, the oblate spheroidal wave functions are appropriate. These functions were not tabulated to an extent sufficient to allow numerical computations when the wavelength of light λ exceeded one tenth the circumference of the disk or aperture. Hence, Meixner and Andrejewski gave only the first term of the power series (applicable to small disks and apertures) for the case of normal incidence (1950); Andrejewski (1953) presented tables permitting calculations for larger disks ($ka = 10$, where $k = 2n/\lambda$ is the wavenumber and a is the radius of the disk). Bouwkamp's remark regarding Meixner and Andrejewski's paper is worth noting: "This new solution is remarkably simple and is to be considered as one of the most outstanding results of electromagnetic theory of diffraction" (Bouwkamp, 1950b).

Other rigorous treatments are by Flammer (1953a, b), who defined and utilized oblate spheroidal vector wave functions, and Nomura and Katsura (1955), who employed an expansion of the hypergeometric polynomial.

Noncircular aperture shapes have not been treated successfully in such a rigorous, analytic manner. As a result various approaches have been tried. When the aperture is small compared to the wavelength of light and one is interested in the far zone behavior, one can use Bethe's concept of associating an effective dipole moment with the aperture. With this technique in mind, Cohn (1951, 1952) obtained values for the electric and magnetic polarizabilities of small apertures of various shapes by electrolytic tank measurements. More recently others have presented numerical results for such polarizabilities (De Meulenaere and Van Bladel, 1977; Van Bladel, 1979; De Smedt and Van Bladel, 1980; Okon and Harrington, 1981; Arvas and Harrington, 1983). Jaggard and Papas (1978) used the concept of symmetrization to define upper and lower bounds on the polarizabilities. And Suzuki (1956) calculated a related quantity, the transmission coefficient, for rectangular apertures assuming different aperture fields.

If one is interested in the near field behavior of the aperture diffraction problem, however, it is necessary to determine the aperture fields. One approach is to solve integral equations for the diffracted aperture field by numerical moment methods. The fundamentals of this approach are described in a book by an early contributor to the field of moment methods (Harrington, 1968). Rahmat-Samii and Mittra (1974)

derived integral equations that are well suited to small apertures. They presented results for the surface currents on a rectangular plate and for the aperture fields in a one wavelength square aperture. In a subsequent paper the same authors constructed numerical results for the aperture field and the on-axis diffracted field of a subwavelength square aperture; they compared their results with those for circular apertures (Rahmat-Samii and Mittra, 1977). Other numerical calculations are available (Mautz and Harrington, 1976; Rao, Wilton, and Glisson, 1982; Sarkar, Arvas, and Rao, 1985).

The work by Kieburtz and Ishimaru (1962) and Kieburtz (1962) is also noted. They computed values for the cross section and the aperture fields in a plane screen containing a doubly periodic array of rectangular apertures. Their variational method of analysis relied on choosing an appropriate trial function for the aperture field components to solve three coupled integral equations.

The behavior of the electric and magnetic fields near aperture corners has attracted recent interest. De Smedt and Van Bladel applied their study of the fields near a metallic cone of arbitrary cross section (1986) to the special case of a flat sector (1987). Using low frequency techniques they expressed the zero-order static potential for the electric field as the product of a radial factor and a spherical harmonic function. Numerical techniques were employed to solve the eigenfunction problem for the spherical harmonics. They presented values for the singularity exponents as a function of the half-maximum opening angle of the flat sector. Their

results have particular relevance as a means for accelerating the accuracy and convergence of numerical algorithms.

Diffraction by polygonal apertures has been of interest for many years. Using the symmetry properties of Fraunhofer diffraction by apertures possessing a symmetry, Komrska (1972) derived a function describing Fraunhofer diffraction by apertures that are regular polygons. He computed diffracted intensity maps for an equilateral triangle, a square, a regular pentagon, a hexagon, and an octagon; he compared them to photographs of the diffraction patterns (Komrska, 1973). A review of earlier research on this topic is contained in part I of Komrska's investigations (Komrska, 1972). Witmer's (1926) paper was overlooked in Komrska's bibliography.

Smith and Marsh (1974) performed similar calculations for a general polygonal aperture by first considering an isosceles triangle and a trapezoid. They obtained the result for a general aperture by considering the aperture to be composed of elemental trapezoids. Witz (1975) presented a different, yet versatile method for obtaining the same result. He evaluated the singularities of a secant-length function, which is determined from the aperture function, and its derivatives. These singularities completely determine the Fourier transform or diffraction pattern.

Sillitto and Sillitto (1975) used a Fourier approach to derive an expression for Fraunhofer diffraction by triangular apertures. The simple expression involved only sinc functions of two nonorthogonal coordinates related to the lengths and directions of two sides of the triangle. Later,

Sillitto (1978) used this result to analyze an arbitrary aperture as the sum of triangles. These formulas have been rederived by means of the Abbe transform (Komrska, 1982) and the Maggi-Rubinowicz transformation (Ganci, 1984).

A different triangular aperture diffraction problem was posed in the book by Stone (1963). The reader is asked to describe the general appearance of the Fraunhofer pattern formed by a narrow slit bent into an equilateral triangle (problem 9-18, page 187). Such an aperture can be viewed as a simple nested triangular aperture; hence, the problem is a precursor to the nested polygon problem considered in this thesis.

Because decomposing a polygonal aperture into elemental triangular apertures leads to a convenient description of the diffraction pattern of a polygon, a detailed study of these elemental triangles is appropriate. One can view these triangular apertures as bounded plane-screen corners of a given angle. A plane-screen corner is formed by the coplanar intersection of two perfectly conducting, vanishingly thick, half-plane screens oriented at the given angle.

One method used in this thesis to evaluate diffraction by such a structure is direct computation of the Fourier transform of a bounded plane-screen corner. Livanos and George (1975) presented closed-form solutions for the optical transform patterns of a semicircular aperture illuminated by a unit amplitude convergent wave and of an edge illuminated by a Gaussian amplitude convergent wave. The relevance to the topics studied in Chapter 3 lies in realizing that an edge is a plane-

screen corner of opening angle 180°. Also, a semicircular aperture is a 180° sector of a circular aperture.

Another method used in this thesis is to expand the diffracted field in Gaussian beam modes. The propagation of Gaussian beams and the expansion of functions in Hermite-Gaussian and Laguerre-Gaussian functions have been studied for many years. Hermite-Gaussian functions are eigenfunctions of the harmonic oscillator Hamiltonian (Liboff, 1980). In the context of optical resonators, Boyd and Gordon (1961) determined the fields of a confocal resonator (rectangular aperture) to be Hermite-Gaussian functions; Boyd and Kogelnik (1962) and Vainshtein (1963) determined the fields of a generalized confocal resonator (circular aperture) to be Laguerre-Gaussian functions. Goubau and Schwering (1961) introduced the concept of electromagnetic wave beams in which the cross-sectional amplitude distribution is repeated at a certain distance from the origin of the beam. They showed that the mode patterns are describable in terms of associated Laguerre-Gaussian functions. Expanding the kernel of the resonator integral equation in a special set of Schmidt functions (Schmidt expansion, see Smithies, 1958) is a method for obtaining the field distributions of the resonator modes (Streifer and Gamo, 1964; Heurtley, 1964; Streifer, 1965; Heurtley and Streifer, 1965); again the Gaussian modes are important. In a comprehensive review article Kogelnik and Li (1966) describe much of the early laser beam and resonator theory. In particular, they sketched how the Gaussian modes are solutions to the paraxial scalar Helmholtz equation and form a complete and orthogonal set of functions.

Gaussian beam expansions have been used for pattern synthesis. Papas (1965) sketched an approximate method due to Bouwkamp and de Bruijn (1946). To solve the integral equation connecting a radiation pattern with the necessary current distribution, first they expanded the radiation pattern in a power series. Then they obtained the current distribution by summing appropriate Hermite-Gaussian functions multiplied by the radiation coefficients and determined factors.

Recently, Bogush and Elkins (1986) developed an approximate theory of large aperture antennas using a Gaussian field expansion. For a separable rectangular geometry, they expanded the aperture field, and hence the diffraction field, in terms of Hermite-Gaussian functions. They presented optimized scaling factors and series coefficients for a number of aperture distributions. For a circular geometry, they introduced an expansion using Laguerre-Gaussian functions that did not allow a polar angular variation in the fields.

Gaussian modes without polar angular variation have also been used to study diffraction of a Gaussian beam through a circular aperture. Tanaka, Saga, and Mizokami (1985) used the first term in a Gaussian beam expansion to investigate the angular divergence of a diffracted Gaussian beam. In a subsequent paper, Tanaka, Yoshida, and Taguchi (1988) used the same approach to approximate the diffraction field through a sequence of three apertures.

We also note an important expansion scheme based on elementary Gaussian beams. We elucidate neither the usefulness nor the difficulties of this expansion, but it is mentioned here for completeness. The Gabor

(1946) representation was introduced within the context of communication theory. After proving that a signal's specificity simultaneously in time and frequency is fundamentally limited by a lower bound on the product of its bandwidth and duration (uncertainty relation for information), Gabor found a general family of signals that achieve the theoretical lower limit of joint uncertainty in time and frequency. These elementary signals, or "logons," are characterized by a Gaussian envelope modulated by a sine wave. The Gabor representation is based on Gabor's suggestion that an expansion in terms of properly shifted elementary signals can represent arbitrary signals. Recently investigators have concentrated on using Gabor's ideas to describe information content (Bastiaans, 1982), aperture theory (Einziger, Raz, and Shapira, 1986), and propagation through a complicated environment (Maciel and Felsen, 1987).

Finally, we point out that an important connection can be made between optical diffraction by apertures and X-ray diffraction by crystals. The book by Harbum, Taylor, and Welberry (1975) states that studying optical analogues of X-ray diffraction phenomena helps in the interpretation of X-ray diffraction patterns by "stimulating thought and in aiding the development of intuition." The book is a collection of optical diffraction pattern photographs of objects ranging from simple apertures to complex combinations of apertures. In particular, plate 23 shows diffraction patterns from helices viewed from different pitches and sampled at equal intervals. The photographs are striking, and one can relate some similarities to the grating-like section patterns in §4.2.1.

1.3 Statement of Notations and Conventions

In this thesis we use the *mks* system of units. The electric permittivity and magnetic permeability of the medium are ε and μ, respectively; the speed of light is c. The time dependence of electromagnetic field quantities is not shown explicitly; we assume it to be $exp(-i\omega t)$. In all cases monochromatic light with wavelength λ is assumed, and the wavenumber is $k = 2\pi/\lambda$. We portray vector quantities as bold face, e.g., $\mathbf{E}$, and unit vectors as $\mathbf{1}_x$ (which specifies a unit vector in the x direction). In Chapter 2 the E_x^A notation means that the field component is evaluated at $z = 0$ and $(x,y) \in A$, where A denotes the aperture, e.g.,

$$E_x(x,y,0) = E_x^A(x,y) \quad for (x,y) \in A \quad .$$

Throughout the thesis we represent the Fourier transform of a function by the same letter but with a bar over it, e.g. $\bar{f}$,

$$\bar{f}(u) = \int_{-\infty}^{\infty} f(x) \, exp(-i2\pi x u) \, dx \quad .$$

We use the following definition for the *sinc* function:

$$sinc(x) = \frac{sin(\pi x)}{\pi x} \quad .$$

We use the term optical intensity to mean the square modulus of the electric field. The term optical transform is taken to mean the optical intensity observed in the focal plane of a converging monochromatic light wave; it is related directly to the square modulus of the Fourier transform of the aperture distribution. In Chapters 3 and 4 we analyze diffraction patterns in a plane perpendicular to the z axis. Sometimes we refer to observation points in this plane by cartesian coordinates (x,y,z) and sometimes by cylindrical coordinates (r,θ,z).

CHAPTER TWO

SMALL APERTURES

In this chapter we present a hypothesized solution for the aperture electric field in a small square aperture ($2a \times 2a$, $ka \ll 1$, where $k = 2\pi/\lambda$ is the wavenumber) in a thin, perfectly conducting plane screen illuminated by a normally incident, linear polarized plane wave of wavelength λ. We outline an integro-differential equation framework and restrict it to small apertures. After hypothesizing the square aperture solution, we substantiate it within the integro-differential equation framework, compare it with the circular aperture solution, and compare it with published numeric results. A final discussion of the far field diffraction pattern makes the usefulness of the results clear.

2.1 Integro-Differential Equation Formulation

Consider a perfectly conducting plane screen of vanishing thickness with aperture A at the plane $z = 0$ as in Fig. 2.1. An electromagnetic field is incident from the left, $z < 0$. We assume an $exp(-i\omega t)$ time dependence. We write the electric field in the presence of the screen as

$$\mathbf{E} = \begin{cases} \mathbf{E}^0 + \mathbf{E}^d & z < 0 \\ \mathbf{E}^d & z \geq 0 \end{cases} \tag{1}$$

where $\mathbf{E}^0$ represents the short-circuit field or the field that would exist if the screen had no aperture, i.e., incident field plus reflected field, and $\mathbf{E}^d$ represents the diffracted field. A similar expression can be written for H. We write the diffracted field as the curl of a vector potential F:

17
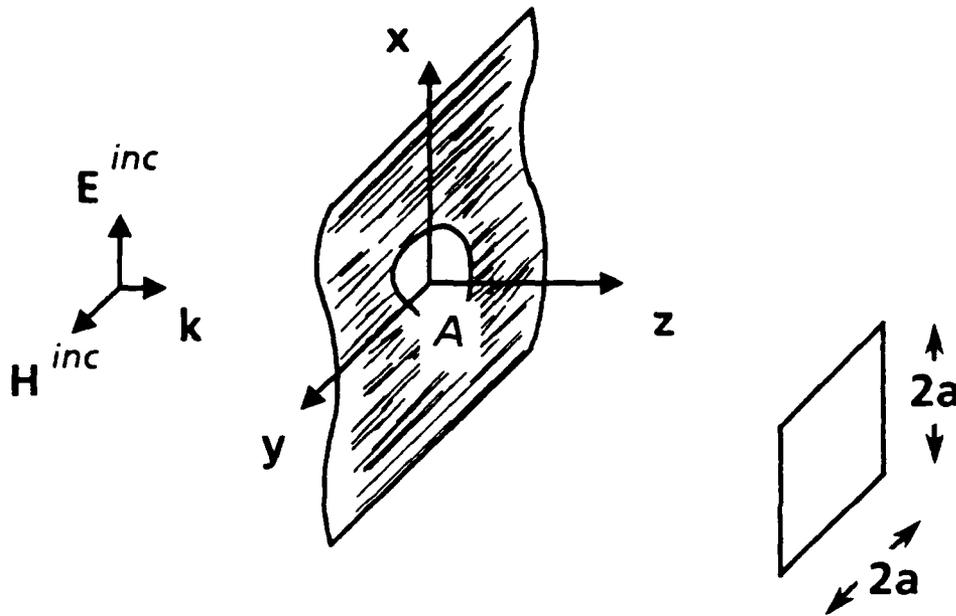
Fig. 2.1. Thin plane screen diffraction geometry. In general an electromagnetic wave is incident from the left $z < 0$. If the screen has an aperture A, then the problem is to find the diffracted field $\mathbf{E}^d$ for $z \geq 0$. Shown here is the specific case of a normally incident plane wave linearly polarized along x. The inset shows the $2a \times 2a$ square aperture considered in the text.

$$\mathbf{E}^d(\mathbf{r}) = \pm \frac{1}{\varepsilon} \nabla \times \mathbf{F}(\mathbf{r}) \ , \tag{2a}$$

where $\mathbf{F}$ is defined by

$$\mathbf{F}(\mathbf{r}) = \frac{\varepsilon}{2\pi} \int \int_A \mathbf{1}_n \times \mathbf{E}^A(\mathbf{r}') \frac{exp\,(ikR)}{R} \, dx' \, dy' \ , \tag{2b}$$

$$R = \left| (x-x')^2 + (y-y')^2 + z^2 \right|^{1/2} \ ,$$

$$k = \frac{\omega}{c} = \frac{2\pi}{\lambda} \ .$$

In Eq. (2a) we choose the plus sign if $z \geq 0$ and the minus sign if $z < 0$. ε is the permittivity of the medium. In Eq. (2b) $\mathbf{1}_n$ is a unit vector normal to the screen pointing in the $+z$ direction, $\mathbf{E}^A(\mathbf{r}')$ is a notation to emphasize that

the value of the electric field in the aperture, $(x',y') \in A$, $z = 0$, is used, and the integration is over the aperture only. This notation to denote $(x',y') \in A$, $z = 0$ will be used throughout this chapter, e.g., $\mathbf{E}(x',y',0) = \mathbf{E}^A(x',y')$. The scattered magnetic field, $\mathbf{H}^d$, can be obtained from Eq. (2a) by means of the appropriate Maxwell curl equation.

It is instructive to pause and consider some of the properties and results of Eqs. (2a) and (2b). First, the tangential components of the electric field in the aperture determine completely the vector potential. Hence, $\mathbf{F}$ has only an x and a y component. The uniqueness theorem for electromagnetic waves in steady state contends that the electromagnetic field within an empty bounded region is determined uniquely by the tangential components of the electric field (or magnetic field) on the boundary surfaces of the region. The right half-space in the diffraction problem is bounded by the infinite screen plus the aperture and an infinitely large hemisphere. The tangential electric field values on the screen and the hemisphere are zero; thus, only the aperture fields are needed to determine the electromagnetic field in the right half-space.

The symmetries of the equations are important, too. It is evident that $\mathbf{F}$ is an even function of z. Consequently (Jackson, 1975, pp. 435–438),

$$E_x^d, E_y^d, H_z^d \quad \text{are even functions of } z \quad,$$

$$H_x^d, H_y^d, E_z^d \quad \text{are odd functions of } z \quad. \tag{3}$$

For example, if $z_0 > 0$,

$$E_x^d(x,y,-z_0) = -\frac{1}{\varepsilon}\left[-\frac{\partial}{\partial z} F_y(x,y,z)\right]_{z=-z_0} \quad,$$

$$E_x^d(x,y,-z_0) = -\frac{1}{\varepsilon}\left[+\frac{\partial}{\partial z} F_y(x,y,z)\right]_{z=+z_0} \quad,$$

$$E_x^d(x,y,-z_0) = E_x^d(x,y,+z_0) \qquad QED$$

Finally, if **F** is rewritten as

$$\mathbf{F}(\mathbf{r}) = \frac{\varepsilon}{4\pi} \int\!\!\int_A \mathbf{M}(\mathbf{r}') \frac{exp(ikR)}{R} \, dx' \, dy' \quad , \qquad (4a)$$

where

$$\mathbf{M}(\mathbf{r}') = 2\,\mathbf{1}_n \times \mathbf{E}^A(\mathbf{r}') \quad , \qquad\qquad (4b)$$

then we recognize **F** as an magnetic vector potential (analogous to **A**, the electric vector potential, i.e., **B** = **∇** × **A**), and **M** is the effective magnetic current density in the hole.

Smythe (1947) used a similar physical insight to derive the expression for E^d. He considered a thin plane, double current sheet in which the distance between layers is very small and the current densities in the two layers at any point of the sheet are equal and opposite. He showed that a double current sheet giving any desired electric field could be built up out of infinitesimal solenoids. The radiation from these small oscillating current loops can be added up to yield the previous equations.

Digressing a bit further, a step-by-step procedure for obtaining Eqs. (1) and (2) is shown in Fig. 2.2. The figure and the succeeding remarks are attributable to Butler et al. (1978). "The original problem is depicted in Fig. 2(a), while in Fig. 2(b) the aperture is short-circuited, i.e., the conducting screen is made continuous, and the electric field is restored to its original value E_t^A at $z = 0^-$ [$z = 0^-$ means the limit as $z \to 0$ from the left] by the equivalent surface magnetic current $\mathbf{M}_s$ (= $\mathbf{1}_z \times \mathbf{E}_t^A$) placed over the region A on the short-circuited screen. From Fig. 2(b) one obtains Fig. 2(c) directly by use of image theory." The fields in the left half-space

21
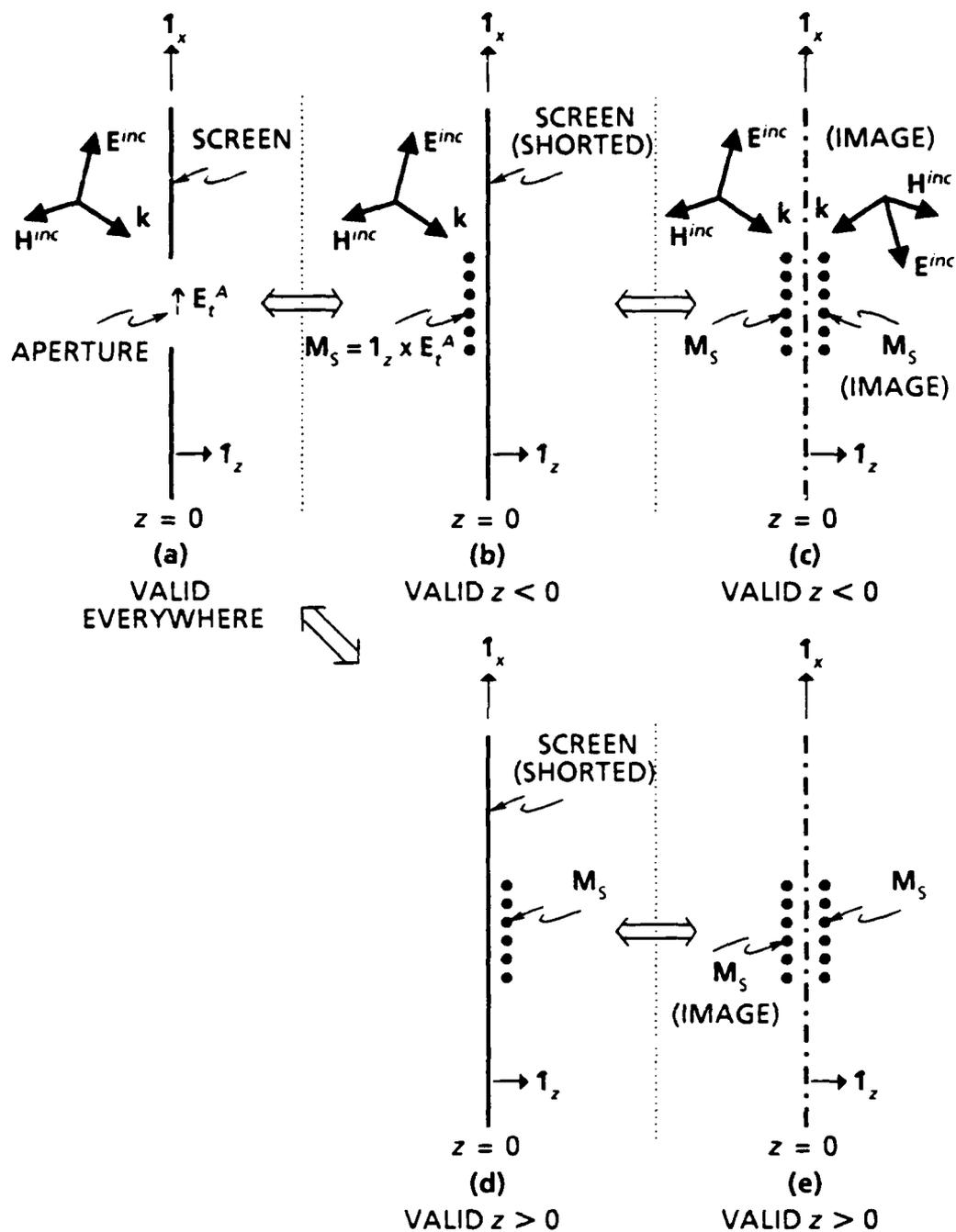
Fig. 2.2. Half-space equivalences. The original problem is depicted in (a).
The left half-space problem is depicted in (b) and (c); the right half-space
problem is depicted in (d) and (e).

can now be written down since the currents, incident field, and image fields reside in a homogeneous space of infinite extent. The short-circuit field E_0 is the sum of the incident field and its image; the diffracted field arises from the equivalent magnetic current plus its image. The vector potential for this latter contribution is F as defined in Eq. (2b). Figures 2(d) and 2(e) illustrate the right half-space problem. In this case the short-circuit field is zero, and the entire radiated field is determined by the equivalent current and its image.

Continuity of the electric field in the hole implies that

$$E_z(0^-) = E_z(0^+) \quad ,$$

$$E_z^0(0^-) + E_z^d(0^-) = E_z^d(0^+) \quad , \tag{5}$$

where $E_z(0^\pm)$ denotes the value of E_z for small values of $\pm|z|$ and $(x,y) \in A$. In consequence, from the symmetry properties (3)

$$E_z^0(0^-) = 2E_z^d(0^+) \quad . \tag{6}$$

From image theory, the normal component of the short-circuit electric field is twice the normal component of the incident electric field. Therefore,

$$E_z(0) = E_z^{inc}(0) \quad . \tag{7}$$

The distinction between 0^+ and 0^- is no longer necessary. Similarly,

$$H_t(0) = H_t^{inc}(0) \quad . \tag{8}$$

where H_t denotes tangential components of H. That is, in the aperture the normal component of E and the tangential components of H are unaffected by the presence of the screen. The physical origin of this statement is rooted in the assumption about an infinitesimally thick screen. The surface current density that is necessarily produced in the screen to satisfy the boundary conditions can have no z component. As a result the

scattered electric (magnetic) field in the aperture arising from this current has *no component normal (tangential) to the screen.*

Finally, these continuity requirements and symmetry properties combined with suitable manipulation of the relations describing **E** and **H** in terms of **F** yield the following set of three differential equations that hold in the aperture:

$$\left[\left(\frac{\partial^2}{\partial x^2} \right) + \left(\frac{\partial^2}{\partial y^2} \right) + k^2 \right] F_x^A = -\varepsilon \frac{\partial}{\partial z} E_y^{inc} \quad , \tag{9a}$$

$$\left[\left(\frac{\partial^2}{\partial x^2} \right) + \left(\frac{\partial^2}{\partial y^2} \right) + k^2 \right] F_y^A = \varepsilon \frac{\partial}{\partial z} E_x^{inc} \quad , \tag{9b}$$

$$\left| \frac{\partial}{\partial x} F_y^A - \frac{\partial}{\partial y} F_x^A \right| = \varepsilon E_z^{inc} \quad . \tag{9c}$$

where

$$\nabla_{xy}^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \quad .$$

Note that the components of the vector potential are evaluated at $(x,y) \in A$, $z = 0$. These differential equations together with the integral relation Eq. (2b) constitute Copson's formulation of the boundary value problem (Copson, 1950, Theorem A).

To obtain a unique and physically important solution, a further condition is required: the behavior at the rim of the aperture must be specified. The edge condition formulated by Meixner (1949, 1972; also Van Bladel, 1964) asserts that the energy density be integrable at the edge. This condition leads to certain conclusions about the behavior of electromagnetic fields near sharp edges. The component of the electric

field perpendicular to the edge becomes infinitely large at the edge as $d^{-1/2}$, where d is the distance from the edge. The component of the electric field parallel to the edge vanishes at the edge as $d^{1/2}$. The behavior of the current density, **M**, can be deduced from these properties. The component of the current density parallel to the edge becomes infinite as $d^{-1/2}$; the component perpendicular to the edge vanishes as $d^{1/2}$. This edge behavior is displayed in Sommerfeld's half-plane diffraction problem (Bouwkamp, 1946) and in all known solutions.

A necessary condition for the integro-diferential equations, stated by Bouwkamp (1953), is that

$$E_s = 0 \qquad on\ the\ rim\ of\ A \ , \tag{10}$$

where E_s denotes the projection of **E** upon the tangent to the rim. According to Bouwkamp, the addition of this condition "does not weaken the value of Copson's theorem. On the contrary, it was meant to be and is in fact a further step towards the practical application of the theorem, especially in the construction of approximate solutions." The condition is satisfied explicitly in the solution for a small circular aperture (to be reviewed later) and in the hypothesized solution for a small square aperture (to be presented later).

2.2 Apl lication to Small Apertures (ka ≪ 1)

Now we restrict attention to apertures whose characteristic dimension, a, is small such that $ka \ll 1$. Also, the incident field is considered to be linearly polarized and normally incident, i.e.,

$$\mathbf{E}^{inc} = \mathbf{1}_x exp(ikz) \qquad \mathbf{H}^{inc} = \mathbf{1}_y exp(ikz)/W_0 \ , \tag{11}$$

where 1_x and 1_y are unit vectors in the x and y directions, respectively, and W_0 is the impedance of the medium (see Fig. 2.1).

Expand all of the factors in Eq. (2b) in a power series in ik (Bouwkamp, 1950; Stevenson, 1953; Kleinman, 1978), substitute into Eqs. (9a)–(9c), and retain terms up to first order. The expansions are

$$\mathbf{F}^A = \mathbf{F}_0^A + ik\,\mathbf{F}_1^A + \cdots \quad ,$$

$$1_n \times \mathbf{E}^A = \Lambda_0 + ik\,\Lambda_1 + \cdots \quad ,$$

$$exp(ikR) = 1 + ikR + \cdots \quad .$$

The zero-order potential is

$$\mathbf{F}_0^A(x,y) = \frac{\iota}{2n} \int\int_A \Lambda_0(\mathbf{r}') \frac{1}{R}\,dx'\,dy' \quad . \tag{12}$$

The three differential equations for $\mathbf{F}_0$ are

$$\nabla_{xy}^2\, F_{0x}^A = 0 \quad , \tag{13a}$$

$$\nabla_{xy}^2\, F_{0y}^A = 0 \quad , \tag{13b}$$

$$\frac{\partial}{\partial x} F_{0y}^A - \frac{\partial}{\partial y} F_{0x}^A = 0 \quad . \tag{13c}$$

By requiring that the charge density be integrable over the aperture, Bouwkamp (1954, p. 77, esp. Eq. 9.12) has shown that for electromagnetic radiation normally incident on a circular aperture

$$\Lambda_0 = 0 \qquad ; hence, \quad \mathbf{F}_0^A = 0 \quad .$$

It is reasonable to believe that this is true for any shaped aperture. In other words, no energy gets through the small hole in the zero order.

Thus, to a first-order approximation in ik, the diffraction problem is described by the following differential equations and integral definition:

$$\nabla_{xy}^2 \ F_{1x}^A = 0 \quad , \tag{14a}$$

$$\nabla_{xy}^2 \ F_{1y}^A = \varepsilon \quad , \tag{14b}$$

$$\frac{\partial}{\partial x} F_{1y}^A - \frac{\partial}{\partial y} F_{1x}^A = 0 \quad . \tag{14c}$$

$$F_1^A(x,y) = \frac{\varepsilon}{2\pi} \int \int_A \Lambda_1(r') \frac{1}{R} dx' dy' \quad , \tag{15}$$

where

$$R = \left[(x-x')^2 + (y-y')^2 \right]^{1/2} \quad . \tag{16}$$

Condition (10) becomes

$$\Lambda_{1n} = 0 \quad on \ the \ rim \ of A \quad , \tag{17}$$

where Λ_{1n} denotes the projection of Λ_1 upon the normal to the rim.

2.3 Review of Circular Aperture Solution

The small circular aperture diffraction problem has been solved within the integro-differential equation framework just described. Bouwkamp (1950) presented the details of that derivation (also Van Bladel, 1964, pp. 404–409); we review the main points here.

The basic procedure is to use knowledge about the edge behavior and symmetry properties of the aperture fields to write a qualitative formula for the current density Λ_1. The exact form of Λ_1 is achieved when the set of three differential equations, Eqs. (14a)–(14c), are employed.

Consider a small circular aperture with radius a centered at $(x,y) = (0,0)$ in a perfectly conducting plane screen of vanishing thickness. The aperture is illuminated by the incident field of Eq. (11). Simple symmetry

shows that E_x^A is an even functions of x' and y'. From Eqs. (14c) and (15) one can then deduce that E_y^A is an odd function of x' and y'. Hence, Λ_x is odd and Λ_y is even. The edge behavior of the current density implies that Λ_ρ must be proportional to $(a^2 - \rho'^2)^{1/2}$ at the edge, and Λ_ϕ must be proportional to $(a^2 - \rho'^2)^{-1/2}$ at the edge. As previously stated, the zero-order term is $F_0^A = \Lambda_0 = 0$. For the first-order term, the qualitative behavior of the current density is

$$\Lambda_{1x}(x',y') = \frac{A\,x'\,y'}{\left(a^2 - x'^2 - y'^2\right)^{1/2}} \tag{18a}$$

$$\Lambda_{1y}(x',y') = \frac{B + Cx'^2 + Dy'^2}{\left(a^2 - x'^2 - y'^2\right)^{1/2}} \ . \tag{18b}$$

The constant A in Eq. (18a) should not be confused with the notation denoting the region of the aperture.

The following integrals are needed for the determination of F_1^A:

$$\iint_A \frac{dx'\,dy'}{\left(a^2 - x'^2 - y'^2\right)^{1/2}\left|(x-x')^2 + (y-y')^2\right|^{1/2}} = \pi^2 \quad ,$$

$$\iint_A \frac{x'\,y'\,dx'\,dy'}{\left(a^2 - x'^2 - y'^2\right)^{1/2}\left|(x-x')^2 + (y-y')^2\right|^{1/2}} = \frac{3}{8}\pi^2\,xy \quad ,$$

$$\iint_A \frac{x'^2\,dx'\,dy'}{\left(a^2 - x'^2 - y'^2\right)^{1/2}\left|(x-x')^2 + (y-y')^2\right|^{1/2}} = \frac{\pi^2}{16}(4a^2 + 5x^2 - y^2) \quad .$$

Straightforward computation shows

$$F^A_{1x}(x,y) = \varepsilon \, \frac{3\pi}{16} \, xy \, A$$

$$F^A_{1y}(x,y) = \varepsilon \left| \frac{\pi}{2} B + \frac{\pi}{32} (4a^2 + 5x^2 - y^2)C + \frac{\pi}{32}(4a^2 + 5y^2 - x^2)D \right| \quad .$$

Substitution into Eq. (14b) reveals that

$$C + D = \frac{4}{\pi} \quad . \tag{19a}$$

Equation (14c) provides an additional constraint:

$$5C - D = 3A \quad . \tag{19b}$$

Because there are four unknowns, two more relations are needed.

These are provided by applying the edge condition (17). The radial

component, $\Lambda_{1\rho}$, which must vanish proportional to $(a^2 - \rho'^2)^{1/2}$ at the

edge, is

$$\Lambda_{1\rho} = \Lambda_{1x} \cos\phi' + \Lambda_{1y} \sin\phi' \quad ,$$

$$\Lambda_{1\rho} = \frac{\sin\phi'}{\left(a^2 - \rho'^2\right)^{1/2}} \left| (B + D\rho'^2) + \rho'^2 \cos^2\phi' (A + C - D) \right| \quad .$$

For the correct edge behavior,

$$B + Da^2 = 0 \quad . \tag{19c}$$

$$A + C - D = 0 \quad . \tag{19d}$$

The solution of Eqs. (19a)–(19d) is

$$A = C = \frac{4}{3\pi} \quad ,$$

$$B = -a^2 \frac{8}{3\pi} \quad ,$$

$$D = \frac{8}{3\pi} \quad .$$

Using the definition of Λ in the first order, we can write

$$1_n \times E^A(x',y') = ik\Lambda_1(x',y')$$

$$1_n \times E^A(x',y') = \frac{4ik}{3\pi} \left\{ \frac{x'y'}{\left(a^2 - x'^2 - y'^2\right)^{1/2}} 1_x + \frac{-2a^2 + x'^2 + 2y'^2}{\left(a^2 - x'^2 - y'^2\right)^{1/2}} 1_y \right\} \quad . \quad (20)$$

The desired solution for E^A is

$$E^A(x',y') = -\frac{4ik}{3\pi} \left\{ \frac{2a^2 - x'^2 - 2y'^2}{\left(a^2 - x'^2 - y'^2\right)^{1/2}} 1_x + \frac{x'y'}{\left(a^2 - x'^2 - y'^2\right)^{1/2}} 1_y \right\} \quad . \quad (21)$$

2.4 Synthesis of Square Aperture Solution

We use a similar approach for the small square aperture diffraction problem. A first-order solution for the tangential components of the electric field in a square aperture will be written based on physical reasoning and intuition. This gives Λ_1. The solution to be analyzed contains three undetermined coefficients. The next step is to calculate F_1^A using this solution. Finally, the undetermined coefficients are computed by requiring that the calculated F_1^A satisfy the differential equation.

The hypothesized solution for the aperture field in a square aperture is synthesized from the solution for an infinite slit in a conducting screen. According to other authors (Sommerfeld, 1954, pp. 273–289; Millar, 1960; Kieburtz and Ishimaru, 1962, pp. 668–669) the series

$$E_x^A(x') = \sum_{n=0}^{\infty} \beta_n \left| a^2 - x'^2 \right|^{n-1/2} \quad , \quad (22a)$$

is sufficiently complete to describe the field within a slit of width $2a$ and infinite in the y direction assuming that the slit is illuminated by a normally incident plane wave polarized along x. Likewise, if the slit is

oriented so that it is infinite in the x direction, then, under the same illumination

$$E_x^A (y') = \sum_{n=1}^{\infty} \beta_n \left| a^2 - y'^2 \right|^{n - 1/2} . \qquad (22b)$$

We notice that Eqs. (22a) and (22b) are alike in form, the only difference arising in the limits of the summation. The term $n = 0$ in Eq. (22a) accounts for the allowed singularity of E_x^A. The term $n = 1$ in Eq. (22b) properly describes the vanishing behavior of the electric field at the edge for this polarization.

The hypothesized solution for E_x^A, the dominant component of the electric field, for a $2a \times 2a$ square aperture is the product of Eqs. (22a) and (22b), namely,

$$E_x^A (x',y') = \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} \beta_{nm} \left| a^2 - x'^2 \right|^{n - 1/2} \left| a^2 - y'^2 \right|^{m - 1/2} . \qquad (23)$$

We retain only the first three terms in Eq. (23).

$$E_x^A (x',y') = \beta_{01} \left| a^2 - x'^2 \right|^{-1/2} \left| a^2 - y'^2 \right|^{1/2}$$

$$+ \beta_{11} \left| a^2 - x'^2 \right|^{1/2} \left| a^2 - y'^2 \right|^{1/2}$$

$$+ \beta_{02} \left| a^2 - x'^2 \right|^{-1/2} \left| a^2 - y'^2 \right|^{3/2} . \qquad (24a)$$

We rewrite this as

$$E_x^A (x',y') = \left(\frac{a^2 - y'^2}{a^2 - x'^2} \right)^{1/2} \left| B + C(x'^2/a^2) + D(y'^2/a^2) \right| (ika) , \qquad (24b)$$

where

$$B = (\beta_{01} + a^2 \beta_{11} + a^2 \beta_{02}) / (ika) ,$$

$$C = -a^2\beta_{11} / (ika) \qquad ,$$

$$D = -a^2\beta_{02} / (ika) \qquad .$$

The explicit factor of *ika* is in anticipation of the pending analysis. Note that the factors of a^2 make the constants B, C, and D dimensionless. Equation (24b) constitutes the hypothesized aperture field solution for the dominant component of the electric field.

As was the case with the circular aperture problem, even though the incident field is polarized along the x axis, the y component of the aperture electric field is nonzero. Nevertheless, it turns out that E_y^A is appreciable only near the edge of the aperture in both cases. To complete the description of the aperture field, a solution for E_y^A is required. Such a solution will not be presented because one can deduce that E_y^A does not contribute to the leading term in the far field diffraction pattern. This statement will be validated in §2.6 by virtue of the following property. For any aperture shape symmetric in x' and y', e.g., circle or square, E_y^A will be an odd function.

2.5 Evaluation of Hypothesized Solution

2.5.1 Computation of F_1^A and $\nabla^2 F_1^A$

The first step in evaluating the validity of the synthesized solution for E_x^A, Eq. (24b), is to calculate F_1^A as defined in Eq. (15). Since only the x component of E^A is being considered, only the y component of Λ_1 and hence of F_1^A need be computed. From Eq. (24b) and the power series expansion of

$$1_n \times E^A$$

$$\Lambda_{1y}(x',y') = (a) \left(\frac{a^2 - y'^2}{a^2 - x'^2} \right)^{1/2} \left| B + C(x'^2/a^2) + D(y'^2/a^2) \right| \quad . \tag{25}$$

Now we write Eq. (25) in normalized coordinates, namely,

$$\xi' \equiv x'/a \quad ; \quad \eta' \equiv y'/a \quad ,$$

$$\Lambda_{1y}(x',y') = (a) \left(\frac{1 - \eta'^2}{1 - \xi'^2} \right)^{1/2} \left| B + C\xi'^2 + D\eta'^2 \right| \quad . \tag{26}$$

Rewriting Eq. (15) in normalized coordinates gives

$$F_{1y}^A(\xi,\eta) = \frac{\varepsilon a^2}{2\pi} \int_{-1}^{1} \int_{-1}^{1} \left(\frac{1 - \eta'^2}{1 - \xi'^2} \right)^{1/2} \left| B + C\xi'^2 + D\eta'^2 \right|$$

$$\times \frac{d\xi' \, d\eta'}{[(\xi - \xi')^2 + (\eta - \eta')^2]^{1/2}} \quad , \tag{27}$$

where

$$\xi \equiv x/a \quad ; \quad \eta \equiv y/a \quad .$$

The differential equation (14b) becomes

$$\nabla_{\xi\eta}^2 F_{1y}^A = \varepsilon a^2 \quad . \tag{28}$$

At this point we desire to evaluate analytically $F_{1y}{}^A$, Eq. (27), for points (ξ,η) in the aperture, and then to substitute that analytic expression into Eq. (28). We have not accomplished this. One recourse is to evaluate the integral expression and differential equation numerically. Before proceeding in this way, though, we transform the ξ' integration into elliptic integrals by standard techniques (see Appendix A: Transformation of Elliptic Integrals into Standard Form). This greatly reduces the numerical computation required to evaluate $F_{1y}{}^A$ at a point (ξ,η) within the aperture. Equation (27) becomes

$$F_{1y}^A(\xi,\eta) = \frac{\varepsilon a^2}{2\pi} \int_{-1}^{1} (1 - \eta'^2)^{1/2} \left| B I_2 + C I_3 + D\eta'^2 I_2 \right| d\eta' \quad , \tag{29}$$

where

$$I_2 = \frac{2}{(\lambda_1 - \lambda_2)^{1/2}} \, K(k) \quad , \tag{30a}$$

$$I_3 = \left(\frac{\xi^2}{1+\lambda_2} \right) \left| \frac{2}{(\lambda_1 - \lambda_2)^{1/2}} \right| \left[\frac{K(k)}{1+\lambda_1} - P(n,k) - \left(\frac{\lambda_2 - \lambda_1}{1+\lambda_1} \right) E(k) \right] \quad , \tag{30b}$$

$$\lambda_{1,2} = \frac{1}{2} \left\{ [\xi^2 + o^2 - 1] \pm [(\xi^2 + o^2 - 1)^2 + 4o^2]^{1/2} \right\} \quad ,$$

$$o^2 = (\eta - \eta')^2 \quad , \quad k^2 = \frac{\lambda_2}{\lambda_2 - \lambda_1} \quad , \quad n = \frac{1 + \lambda_2}{\lambda_2 - \lambda_1} \quad .$$

We have used the following definitions for the complete elliptic integrals of the first, second, and third kind (Gradshteyn and Ryzhik, 1980, pp. 904–905):

$$K(k) = F(\frac{\pi}{2}, k) = \int_0^{\pi/2} \frac{d\phi}{(1 - k^2 \sin^2 \phi)^{1/2}} \quad ,$$

$$E(k) = E(\frac{\pi}{2}, k) = \int_0^{\pi/2} (1 - k^2 \sin^2 \phi)^{1/2} \, d\phi \quad ,$$

$$P(n,k) = \Pi(\frac{\pi}{2}, n, k) = \int_0^{\pi/2} \frac{d\phi}{(1 - n \sin^2 \phi)(1 - k^2 \sin^2 \phi)^{1/2}} \quad .$$

Since the constants B, C, and D are not yet known, we must compute each term in Eq. (29) individually. Thus, define

$$F_B^A (\xi, \eta) = \frac{1}{2\pi} \int_{-1}^{1} (1 - \eta'^2)^{1/2} \quad I_2 \, d\eta' \quad , \tag{31a}$$

$$F_C^A(\xi,\eta) \;=\; \frac{1}{2\pi} \int_{-1}^{1} (1 - \eta'^{\,2})^{1/2} \quad I_3 \, d\eta' \qquad , \tag{31b}$$

$$F_D^A(\xi,\eta) \;=\; \frac{1}{2\pi} \int_{-1}^{1} (1 - \eta'^{\,2})^{1/2} \, \eta'^{\,2} I_2 \, d\eta' \qquad , \tag{31c}$$

such that

$$F_{1y}^A = \varepsilon a^2 (B\,F_B^A + C\,F_C^A + D\,F_D^A) = \varepsilon a^2 F_{BCD}^A \; . \tag{31d}$$

As a result Eq. (28) becomes

$$\nabla_{\xi\eta}^{2} F_{BCD}^A = 1 \; . \tag{32}$$

We evaluated the three terms, Eqs. (31a)–(31c), numerically at 10,000 sample points in the aperture. Owing to the symmetry of the problem it was necessary to consider points in one quadrant only. The sample points were equally spaced on a grid such that $0 \le \xi,\eta \le 0.99$; thus, the sample spacing was 0.01. We discuss some details of the numerical calculations in Appendix B: Numerical Integration and Differentiation. Equations (31a)–(31c) are plotted as functions of (ξ,η) in Fig. 2.3. For clarity of presentation, only sample points on a 0.05 grid spacing are shown. In other words, $0 \le \xi,\eta \le 0.95$ with a sample spacing of 0.05.

We then used the numerical results to calculate the left-hand side of Eq. (32) for each term, i.e., $\nabla_{\xi\eta}^{2} (F_B^A, F_C^A, \text{or } F_D^A)$. Values were computed for each term at 400 sample points on the same grid spacing as in Fig. 2.3. Great caution must be exercised in doing these calculations because numerical differentiation is inherently unstable (see Appendix B). The results for each of the three terms are presented in Fig. 2.4.

One might attempt to perform the differentiation indicated in Eq. (32) before carrying out the integrations in Eqs. (31a)–(31c). This makes the integral divergent, however, since points within the aperture are being

35
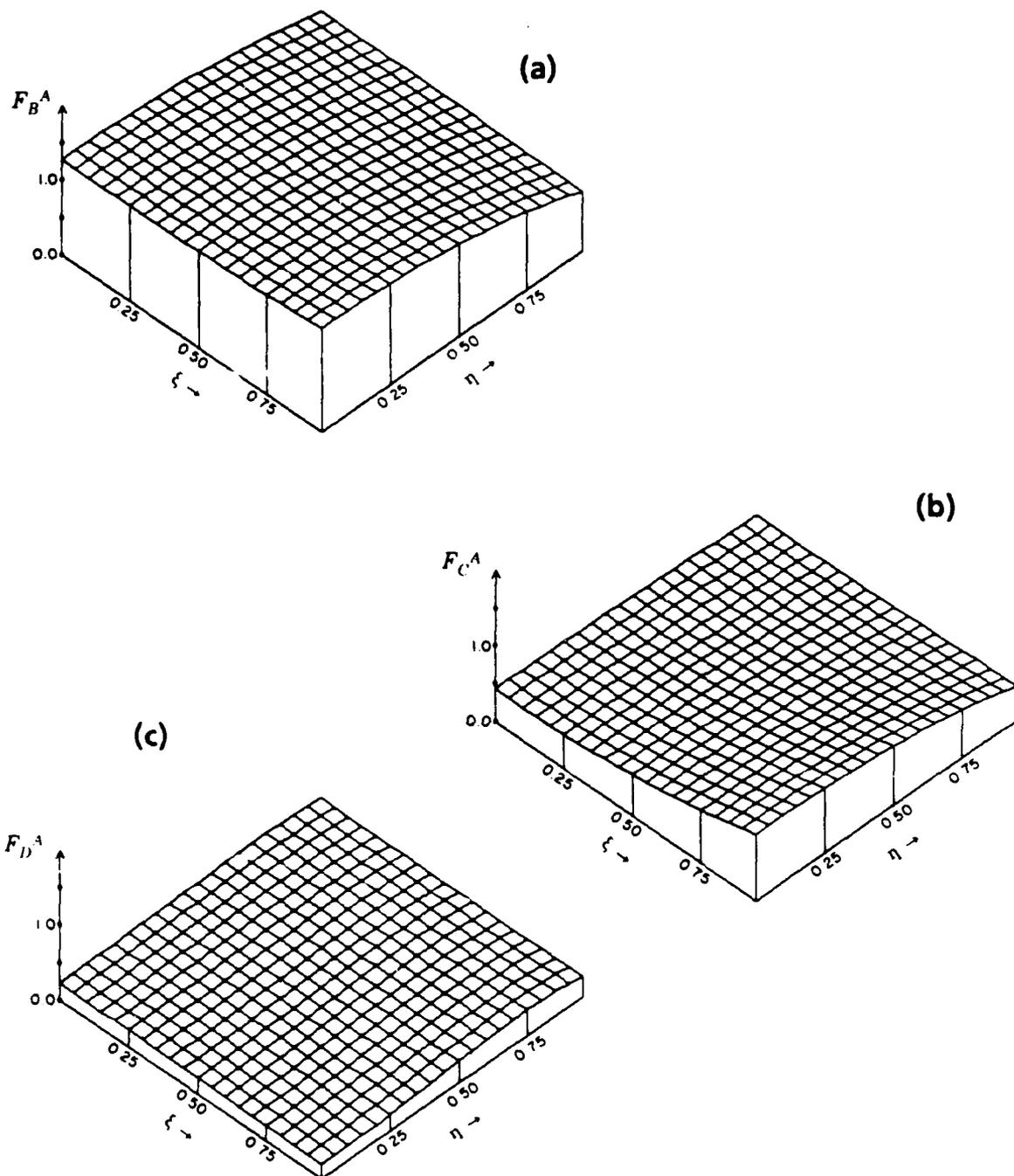
Fig. 2.3. Plots of (a) $F_B{}^A$, (b) $F_C{}^A$, and (c) $F_D{}^A$ over the first quadrant of the aperture. Normalized coordinates are used: $\xi \equiv x/a$, $\eta \equiv y/a$. $0 \le \xi, \eta \le 0.95$.
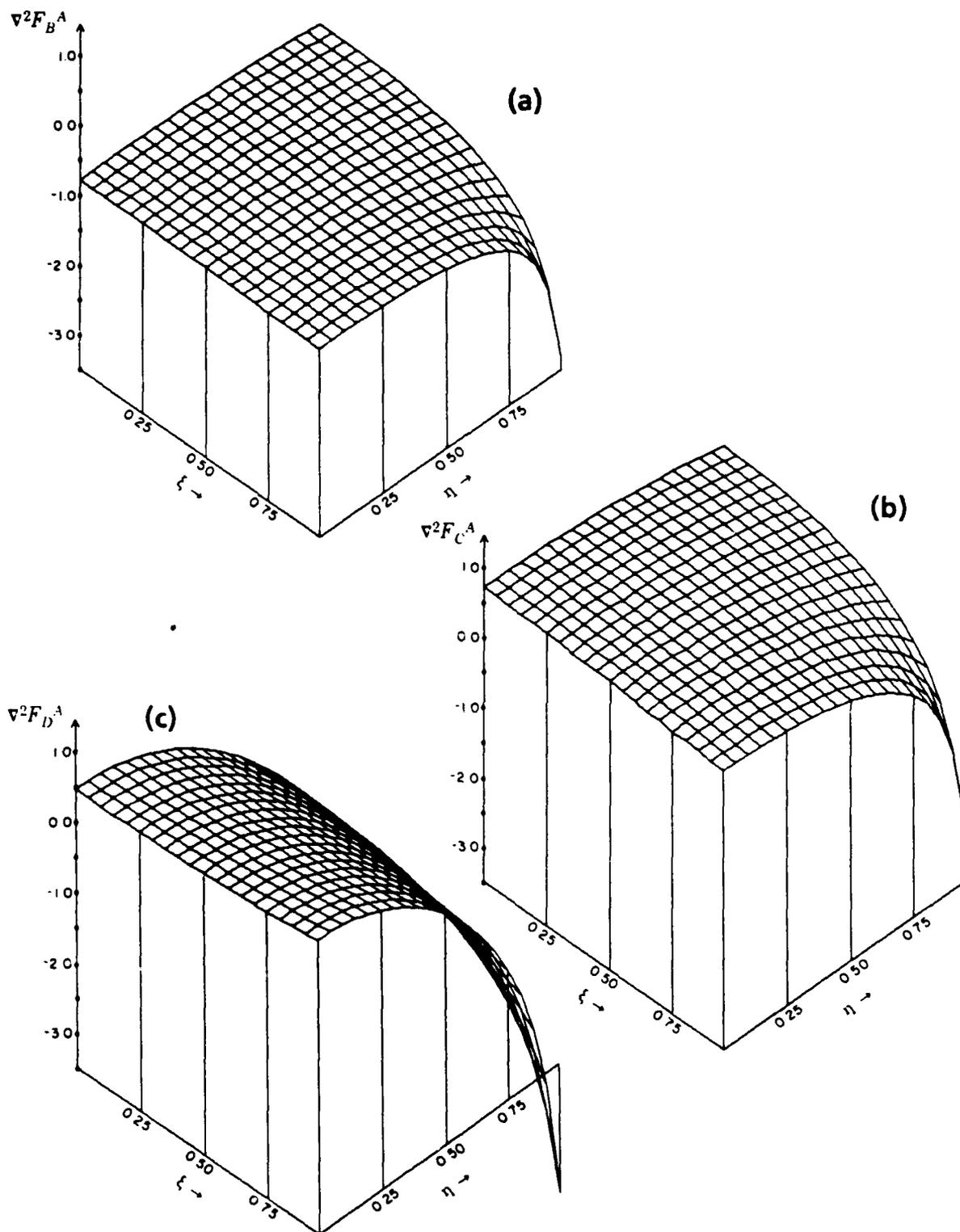
Fig. 2.4. Plots of (a) $\nabla_{\xi\eta}^2 F_B^A$, (b) $\nabla_{\xi\eta}^2 F_C^A$, and (c) $\nabla_{\xi\eta}^2 F_D^A$ over the first quadrant of the aperture. These three surfaces are derived from the corresponding surfaces in Fig. 2.3.

evaluated. This change in the nature of the integral can be seen by computing the contribution from a small circle or radius δ about the field point (ξ, η). The contribution is finite for the integrals in Eqs. (31a)–(31c). But the contribution is infinite with vanishingly small δ if the transverse Laplacian operator in Eq. (32) is moved inside the integral.

2.5.2   Determination of the Unknown Coefficients

The final step in the numerical analysis is to determine the constants B, C, and D such that Eq. (32) is satisfied. There are 400 data points for each of the three curves in Fig. 2.4. These data in conjunction with Eqs. (31d) and (32) can be used to write an overdetermined system of equations for the three unknowns. We solved for these constants by a linear least-squares method and obtained the result

$$B = -0.854 \quad ,$$
$$C = 0.543 \quad , \tag{33}$$
$$D = -0.004 \quad .$$

These results agree with what one might suspect by simply looking at Fig. 2.4. The goal is to obtain a planar surface by adding the three surfaces together, i.e., $\nabla_{\xi\eta}^2 F_{BCD}{}^A = 1$. One can see that the B and C terms (which are generally planar, except in the corners where they are both peaked) will offset each other overall; on the other hand the D term does not seem to have any characteristics that will aid in obtaining a planar surface. A physical argument justifying $D \approx 0$ is presented in the next section.

2.5.3   Discussion of Hypothesized Solution

Using Eq. (24b) and the numerical values for the coefficients (33), we can write a concise mathematical description for the dominant component of the diffracted electric field in a small square aperture.

$$E_x^A(x',y') = (ika)\left(\frac{a^2 - y'^2}{a^2 - x'^2}\right)^{1/2} \left|(-.854) + (.543)(x'^2/a^2) + (-.004)(y'^2/a^2)\right| \quad (34)$$

This function, normalized by a factor of ika, is plotted in Fig. 2.5. Notice that Eq. (34) allows one to evaluate easily the field at any point within the aperture. Further, the nature of the edge behavior of the electric field is displayed explicitly and recognized readily in the formula and in Fig. 2.5. For a fixed value of y, the electric field has a singular behavior near the edges at $x = \pm a$. This singular behavior is precisely the behavior required
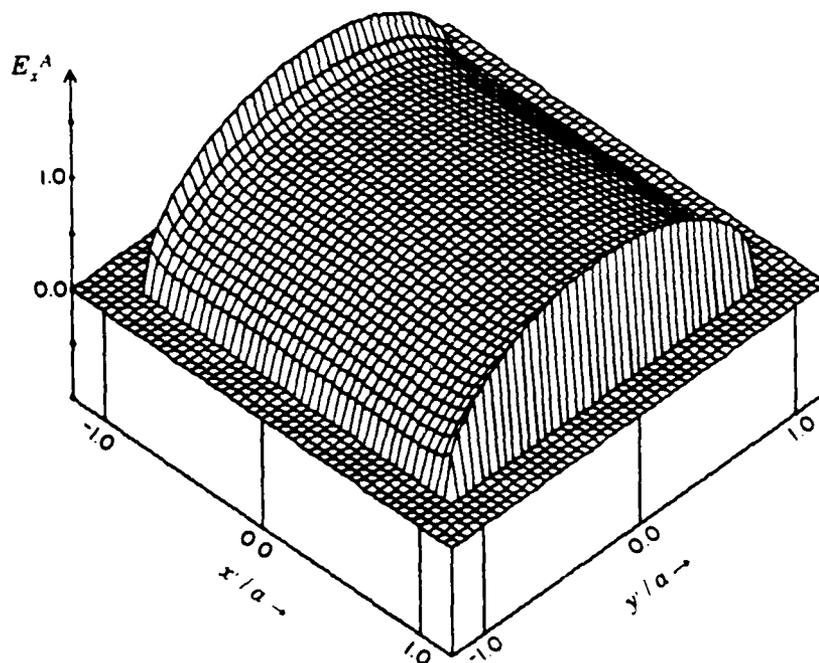
Fig. 2.5. Plot of synthesized solution for E_x^A over the entire aperture. The plot is normalized by ika and is computed from Eq. (34). Outside the aperture, $E_x^A = 0$.

by Meixner's edge condition [see Eq. (10)]. For a fixed value of x, the electric field approaches zero near the edges at $y = \pm a$. The behavior of E_x^A is consisistent with the known behavior of the electric field near asharp edge.

It is appropriate to make a few comments about the behavior of the electric field in the near vicinity of the corners of the aperture. The exact corner behavior of the electric field is not well understood. In order to evaluate fully any proposed solution to this diffraction problem, we must know this behavior and express it as a necessary condition on the validity of the solution as we have done for the edge condition [see Eq. (10)].
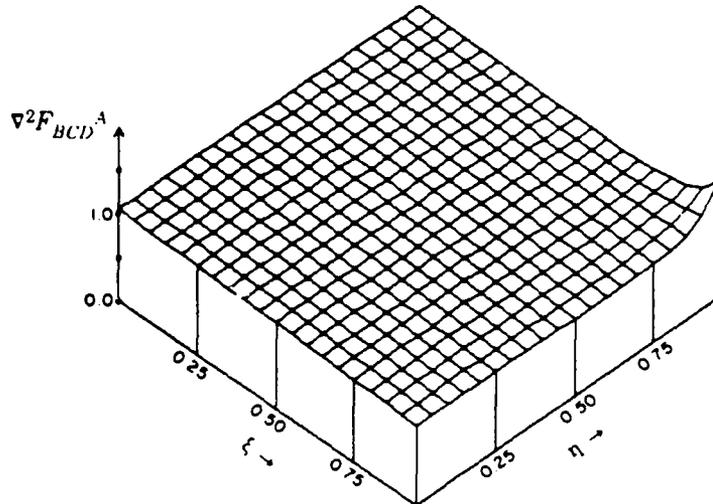
Fig. 2.6. Plot of $\nabla_{\xi\eta}^2 F_{BCD}^A$. The plot is computed by the linear combination of $\nabla_{\xi\eta}^2 F_B^A$, $\nabla_{\xi\eta}^2 F_C^A$, and $\nabla_{\xi\eta}^2 F_D^A$ indicated in Eq. (31d) using Fig. 2.4 and the values for the constants given in (33). This surface shows how well the least-squares fit of the data to Eq. (32) worked. Note that the surface is planar over most of the quadrant; the only deviation is in the close vicinity of the corner, i.e., $(\xi,\eta) = (1,1)$.

Table 2.1. Numerical Histogram of Analysis of
Hypothesized Solution.

Range of $\nabla_{\xi\eta}^2 F_{BCD}{}^A$	# values (out of 400)
(0.90,0.95)	40
(0.95,1.05)	353
(1.05,1.10)	5
(1.10,1.37)	2

Nevertheless, using Eq. (32), we can deduce that the behavior of the electric field in the near vicinity of the corners of the aperture is not represented well by the hypothesized solution. In Fig. 2.6 is plotted the linear combination of $\nabla_{\xi\eta}^2 F_B{}^A$, $\nabla_{\xi\eta}^2 F_C{}^A$, and $\nabla_{\xi\eta}^2 F_D{}^A$ indicated in Eq. (31d) using Fig. 2.4 and the values for the coefficients given in (33). Note that , the vertical axis is labeled $\nabla_{\xi\eta}^2 F_{BCD}{}^A$. These results represent the left-hand side of Eq. (32). This differential equation is satisfied if a value of 1.000 is obtained at every point in the aperture. The values plotted in Fig. 2.6 have an average of 0.998 and a standard deviation of 0.039. The agreement is very good, hence Eq. (34) is validated, except in the close vicinity of the corner, i.e., $(\xi,\eta) = (1,1)$. Table 2.1 shows a breakdown of how many values are in the ranges indicated.

A comparison of Eq. (34) with the solution for a circular aperture of radius a presented earlier is appropriate. The x component of $\mathbf{E}^A$ from Eq. (21) can be written

$$E_x^{A:\ circle}(x',y') = (ika)\left(-\frac{4}{3\pi}\right)\frac{a\,[2 - (x'^2/a^2) - 2\,(y'^2/a^2)]}{(a^2 - x'^2 - y'^2)^{1/2}} \tag{35}$$

The symmetry of both geometries leads one to expect that Eqs. (34) and (35) should have similar behavior along the principal axes, i.e., $x' = 0$ and

$y' = 0$. This is indeed the case.

$$\frac{E_x^{A:\ square}(0,y')}{E_x^{A:\ circle}(0,y')} = \frac{(-0.854) + (-0.004)y'^2/a^2}{(-0.849)} \quad , \tag{36a}$$

$$\frac{E_x^{A:\ square}(x',0)}{E_x^{A:\ circle}(x',0)} = \frac{(-0.854) + (0.543)x'^2/a^2}{(-0.849) + (0.424)x'^2/a^2} \quad . \tag{36b}$$

Equation (36a) is very nearly equal to unity for $|y'| < a$. In addition the single term in the denominator confirms the numerical value for D in (33), i.e., $D = -0.004 \approx 0$. The variance of Eq. (36b) from unity is greater, but the agreement is still relatively good. Such similarity is also apparent in the numerical solutions constructed by Rahmat-Samii and Mittra (1977).

2.6 The Far Field Diffraction Pattern

When the aperture is small compared to a wavelength of light and the observation point is far from the aperture, a multipole expansion of the field is useful. One substitutes Eq. (2b) into Eq. (2a) and approximates R by its power series expansion. We understand the practicality of this expansion when the terms are grouped in a physically meaningful way, i.e., in terms of dipoles, quadrupoles, and so on. The result due to Van Bladel (1977, 1979) is

$$\mathbf{E}^d(\mathbf{r}) = \frac{1}{4\pi} \frac{exp(ikr)}{r} \left\{ -k^2 c\, 1_r \times \mathbf{m}_1 + k^2 1_r \times 1_r \times \mathbf{p} + \frac{ik^3 c}{2} 1_r \times 1_r \bullet \mathbf{Q}_m \right.$$

$$\left. + ik^3 1_r \times \mathbf{m}_2 + \frac{ik^3}{2} 1_r \times 1_r \times 1_r \bullet \mathbf{Q}_e + ... \right\} \quad , \tag{36}$$

where

$$m_1 = \frac{1}{i\omega} \int\int_A M \, ds' \qquad \textit{(first – order magnetic dipole)}, \qquad (37a)$$

$$p = \frac{1}{2} \int\int_A M \times r' \, ds' \qquad \textit{(electric dipole)}, \qquad (37b)$$

$$m_2 = \frac{1}{6} \int\int_A M \times r' \times r' \, ds' \qquad \textit{(second – order magnetic dipole)}, \qquad (37c)$$

$$Q_m = \frac{1}{i\omega} \int\int_A (Mr' + r'M) \, ds' \qquad \textit{(magnetic quadrupole)}, \qquad (37d)$$

$$Q_e = \frac{1}{3} \int\int_A (M \times r'r' + r'M \times r') \, ds' \qquad \textit{(electric quadrupole)}, \qquad (37e)$$

$$M = M(x',y') = 2 \, 1_n \times E^A(x',y') \quad ,$$

$$1_r = r / r$$

(The definitions of the magnetic multipoles used above are formally analogous with the traditional definitions of the electric multipoles.)

The utility of this expansion in treating far field radiation from small apertures is its simplicity. Once the multipole moments are known, we can calculate easily the radiated field. In the treatment here, we will assume that the aperture is small enough so that only the first-order dipole terms need to be considered. Careful examination of these two terms, Eqs. (37a) and (37b), reveal that the equivalent magnetic dipole moment for the small aperture is *in* the plane of the aperture; the equivalent electric dipole moment is directed *normal* to the aperture. The far field radiation pattern is given simply by the fields attributable to these equivalent dipoles.

The physical origin of these dipole terms is the short-circuit fields in the aperture. In other words we assume that the aperture is so small that the fields far from the aperture are determined by the fields that would exist if the aperture were absent, namely E^0 and H^0. Hence, for a perfectly conducting surface, the normal electric field E^0 induces the electric dipole moment, and the tangential magnetic field H^0 induces the magnetic dipole moment. In the problems under consideration, the exciting field is normally incident. Therefore, $E^0 = 0$, and there is no effective electric dipole moment. For the remainder of this section, we consider only the magnetic dipole moment.

It is now appropriate to validate an earlier claim: $E_y{}^A$ does not contribute to the leading term in the far field pattern. To see this we rewrite Eq. (27a) so that the contribution from $E_y{}^A$ is clearly seen.

$$\mathbf{m}_1 = \frac{1}{i\omega}\left[1_x \int\int_A -E_y^A(x',y')\,dx'\,dy' + 1_y \int\int_A E_x^A(x',y')\,dx'\,dy' \right] \quad . \tag{38}$$

When considering symmetric apertures, $E_y{}^A$ is an odd function of x' and y'. Integration of an odd function over symmetric limits gives zero; hence, the contribution of $E_y{}^A$ to Eq. (38) is zero.

$$\mathbf{m}_1 = \frac{1}{i\omega}\cdot 1_y \int\int_A E_x^A(x',y')\,dx'\,dy' \quad . \tag{39}$$

It is now straightforward to compute the effective magnetic dipole moment for the small circular aperture and the small square aperture.

$$\mathbf{m}_1 = -m\,1_y \quad ,$$

$$m^{circle} = \frac{16a^3}{3c} \quad , \tag{40}$$

$$m^{square} = -\frac{[B + C/2 + D/4]n^2 a^3}{c} = \frac{5.759\,a^3}{c} \quad , \tag{41}$$

where c is the speed of light.

Cohn (1951, 1952) measured the polarizabilities of small apertures using electrolytic tank measurements. From his values, we can ascertain the strength of the effective magnetic dipole moment. His value for the square is

$$m_{COHN}^{square} = \frac{8.2880\,a^3}{c} \tag{42}$$

The value obtained in Eq. (41) is about 30% low. This is another way that the inadequacy of Eq. (34) in representing the field in the corners is evident. The value in Eq. (41) is very close to the value in Eq. (40) corresponding to a circle inscribed within the square. If we retain the next higher order term in Eq. (23) [$n = 1$, $m = 2$, so that a fourth term arises in Eq. (24b), namely, $(x'^2 y'^2 / a^4)$], then the result is $m^{square} = 7.264\,a^3 / c$, which is only about 12% low. Alternately, if we acknowledges the corner discrepancy problem before performing the previously described least squares fit to obtain the unknown coefficients (in other words we use only the numerical data from the central region of the aperture), then we obtain a value of $m^{square} = 8.320\,a^3 / c$.

It is clear from this analysis that restricting the general expansion to only the first three terms does not describe accurately the dominant component of the electric field in the close vicinity of the corners. Higher order terms from the general expansion for the aperture field may improve the result. The procedure for doing this is the same as previously described: 1) using the required number of terms from Eq. (23), integrate

analytically one dimension of $F_1{}^A$, [Eq. (31a)]; 2) integrate numerically the other dimension of $F_1{}^A$, [Fig. 2.3(a)]; 3) compute $\nabla_{\xi\eta}{}^2 F_1{}^A$, [Fig. 2.4(a)]; and 4) determine the associated coefficients by least squares.

2.7 Extensions and Generalizations

Rewriting the general hypothesized solution for $E_x{}^A$, Eq. (23), in another form offers notable insight and understanding of the expansion. When we factor out the edge behavior from Eq. (23), the resulting summation is over polynomials of x' and y', i.e.,

$$E_x^A(x',y') = \left(\frac{a^2-y'^2}{a^2-x'^2}\right)^{1/2} \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} \beta_{nm} (a^2-x'^2)^n (a^2-y'^2)^{m-1} \quad . \tag{43}$$

These polynomials are describable as linear combinations of Chebyshev polynomials of the First (T_n) and Second (U_m) kind. In terms of these polynomials, Eq. (43) becomes

$$E_x^A(x',y') = (ika) \left(\frac{a^2-y'^2}{a^2-x'^2}\right)^{1/2} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \gamma_{nm}^{(x)} T_n(x'/a) U_m(y'/a) \quad . \tag{44}$$

The factor of ika is in keeping with the order of the expansion used in the earlier analysis. The modified coefficients, $\gamma_{nm}{}^{(x)}$, are linear combinations of the β_{nm} and are dimensionless. Table 2.2 gives the first few Chebyshev polynomials. Owing to the symmetry of $E_x{}^A$ (even in x' and y'), the summation in Eq. (44) is effectively over only the even values of n and m, i.e.,

$$\gamma_{nm}^{(x)} = 0 \qquad for \ \ n = odd \ \ or \ \ m = odd \quad . \tag{45}$$

The choice of Chebyshev polynomials of the First and Second kind to describe Eq. (43) instead of a different complete set of polynomials is motivated by the leading factor that reflects the edge behavior of $E_x{}^A$. This

Table 2.2. Chebyshev Polynomials of the First and
Second Kind.

n	$T_n(\xi)$	$U_n(\xi)$
0	1	1
1	ξ	2ξ
2	$2\xi^2 - 1$	$4\xi^2 - 1$
3	$4\xi^3 - 3\xi$	$8\xi^3 - 4\xi$

factor relates directly to the orthogonality relations for the Chebyshev polynomials given below.

$$\int_{-1}^{1} \frac{1}{(1-\xi^2)^{1/2}} T_\nu(\xi) T_\mu(\xi)\, d\xi = \begin{cases} \mathrm{n} & \nu = \mu = 0 \\ \mathrm{n}/2 & \nu = \mu \neq 0 \\ 0 & \nu \neq \mu \end{cases} \quad , \qquad (46a)$$

$$\int_{-1}^{1} (1-\xi^2)^{1/2} U_\nu(\xi) U_\mu(\xi)\, d\xi = \begin{cases} \mathrm{n}/2 & \nu = \mu \\ 0 & \nu \neq \mu \end{cases} \qquad (46b)$$

We gain an understanding of the hypothesized expansion if Eq. (44) is used to compute the multipole moments described in §2.6. Because of the orthogonality relations (44), the computation of these multipoles is straightforward. An especially simple, yet important, case is the first order magnetic dipole moment. Inserting Eq. (44) into Eq. (39) and using the orthogonality relations (44) yield

$$m^{square} = -\frac{a^3}{c} Y_{00}^{(x)} \qquad (47)$$

We need only the first term in the Chebyshev expansion for E_x^A to produce the correct far field, i.e., a dipole field. Such a description has a significant advantage over the description of the far field in terms of a point dipole moment: the near field behavior of the electric field expansion

is nonsingular. In contrast, when we attribute the radiation pattern of a small aperture to an effective point dipole at the aperture, we must be careful in applying the approximation close to the aperture. The dipole moment method predicts an infinite field at the aperture. The present expansion does not; yet, the expansion, Eq. (44), is computationally simple and physically significant.

Further, we can make a hypothesis for E_y^A.

$$E_y^A(x',y') = (ika)\left(\frac{a^2-x'^2}{a^2-y'^2}\right)^{1/2} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \gamma_{nm}^{(y)} T_n(y'/a) U_m(x'/a) \quad . \tag{48}$$

The representation is similar to Eq. (44) except that the role of the variables x' and y' have switched. Once again, symmetry will reduce the number of terms in the expansion. Because E_y^A is odd symmetric,

$$\gamma_{nm}^{(y)} = 0 \quad \text{for } n = \text{even or } m = \text{even} \quad . \tag{49}$$

To summarize the hypothesized solution for the electric field in a small square aperture can be used to describe accurately the far field pattern radiated by the aperture. The expansion improves on the simple dipole model for this diffraction problem by permitting evaluation of the diffracted fields at distances closer to the aperture without exhibiting the singular behavior of the dipole method. An additional benefit results from the expansion's explicit description of the edge behavior for the electric field components in the aperture without introducing computational complexity.

CHAPTER THREE
OPTICAL TRANSFORM OF A PLANE-SCREEN CORNER OF
ARBITRARY ANGLE

A plane-screen corner is illustrated in Fig. 3.1. We assume that the diffracting screen, positioned at $z = 0$, is perfectly conducting. The opening angle of the corner θ_0 is symmetric about the x axis. In this chapter we treat this infinite area corner aperture bounded in the following ways: Gaussian illumination (§3.1.2.2), uniform illumination with a slit boundary (§3.2.1), and uniform illumination with a circular boundary (§3.2.2).

The solutions derived in this chapter are achieved within the scope of the Fresnel approximation. We begin with the exact vector integral representation for the diffraction problem, and we describe subsequently a boundary value problem for one component of the electric field. We make familiar approximations to arrive at the Fresnel diffraction integral. A slight modification reveals a second underlying boundary value problem. The Gaussian beam expansion (GBE) arises naturally from this boundary value problem and is a suitable technique for solving the Gaussian illuminated plane-screen corner diffraction problem. After a demonstration and discussion of the GBE solution for the uniform circular aperture diffraction problem, we present Fourier plane diffraction calculations and computer plots for corners of various angles.

We treat the finite corner of §3.2.1 by determining the Fourier transform of the transmission function. The solution is examined carefully
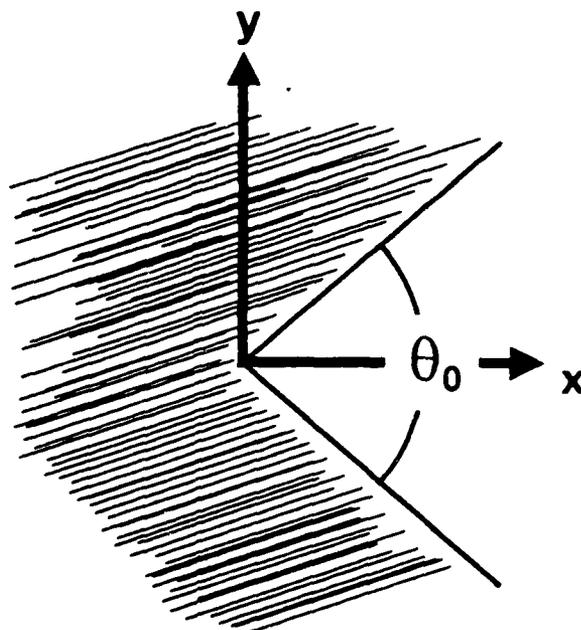
Fig. 3.1. Plane-screen corner of arbitrary angle. The thin, perfectly conducting diffracting screen is positioned at $z = 0$. The opening angle θ_0 is symmetric about the x axis so that the corner can be described mathematically by $rect(\theta/\theta_0)$. The depicted aperture has infinite area. In the chapter the aperture is bounded by a Gaussian (§3.1.2.2), a slit (§3.2.1), and a circle (§3.2.2).

to ensure proper behavior at all points in the Fourier plane. The analysis is straightforward and forms the framework for studying polygonal and nested polygonal apertures in Chapter 4.

Finally, we investigate a plane-screen corner bounded by a circular aperture. The Fourier transform is obtained by explicit integration with respect to one dimension and then expanding the integral in terms of Bessel functions; term-by-term integration of this expansion yields a result that is suitable for computer calculation.

3.1 Gaussian Beam Expansion Analysis

3.1.1 Mathematical Development and Physical Basis

The diffraction geometry is illustrated in Fig. 3.2. The x',y' plane at z = 0 contains a perfectly conducting plane screen with an aperture A. The tangential components of the exact electric field in the aperture specify uniquely the electric field in the right half-space, i.e., $z \geq 0$ [Eqs. (2.2a) and (2.2b)]:

$$E(r) = \frac{1}{2\pi} \nabla \times \int \int_A 1_n \times E(r') \frac{exp(ikR)}{R} dx'dy' \quad , \qquad (1)$$

where

$$R \equiv |r - r'| = \left[(x-x')^2 + (y-y')^2 + z^2 \right]^{1/2} \quad ,$$

$exp(ikR)/R$ is the free-space Green's function, 1_n is a unit normal pointing
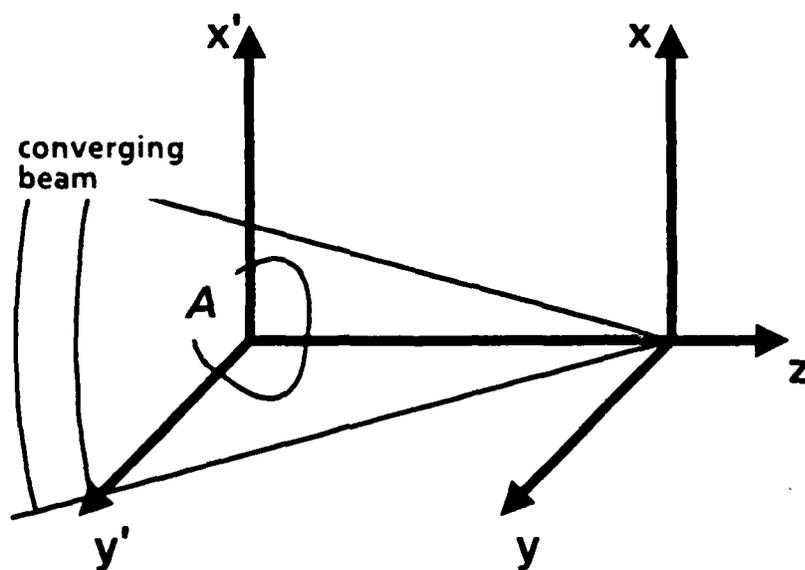
Fig. 3.2. Geometry for diffraction problems. The x',y' plane, i.e., z = 0, contains a thin, perfectly conducting plane screen with an aperture A. One seeks to determine the diffraction pattern in the x,y plane at a distance z from the aperture. The figure shows a converging beam illuminating the aperture and observation of the diffraction pattern in the focal plane, but this is illustrative and not restrictive.

in the $+z$ direction, and $E(r')$ is the exact electric field in the aperture. The integration extends over only the aperture because the screen is a perfect conductor.

As an approximation to the exact electric field in the aperture, assume that $E(r')$ may be replaced by $E^{inc}(r')$. This substitution in Eq. (1) yields

$$E(r) = \frac{1}{2n} \nabla \times \int\int_A 1_n \times E^{inc}(r') \frac{exp(ikR)}{R} dx'dy' \quad . \tag{2}$$

This approximation is reasonable for apertures that are much larger than a wavelength.

It is useful to write out explicitly the expression for all three components of the electric field, noting that the del operator in Eq. (2) acts on the coordinates of the observation point (x,y,z):

$$E_x(x,y,z) = \frac{1}{2n} \int\int_A E_x^{inc}(x',y',0) \frac{exp(ikR)}{R} \left(ik - \frac{1}{R} \right)\left(-\frac{z}{R} \right) dx'dy' \quad , \tag{3}$$

$$E_y(x,y,z) = \frac{1}{2n} \int\int_A E_y^{inc}(x',y',0) \frac{exp(ikR)}{R} \left(ik - \frac{1}{R} \right)\left(-\frac{z}{R} \right) dx'dy' \quad , \tag{4}$$

$$E_z(x,y,z) = \frac{1}{2n} \int\int_A \left[\left(\frac{x-x'}{R} \right) E_x^{inc}(x',y',0) + \left(\frac{y-y'}{R} \right) E_y^{inc}(x',y',0) \right]$$

$$\times \frac{exp(ikR)}{R} \left(ik - \frac{1}{R} \right) dx'dy' \quad . \tag{5}$$

In Eqs. (3)–(5) we see that the electric field at any point in the right half-space is specified completely by the tangential components of the electric field in the aperture. Since we have required implicitly the fulfillment of Sommerfeld's radiation condition, this follows from the uniqueness theorem for Maxwell's equations in the harmonic time-dependent case, too.

Now consider the incident electric field to be a normally incident electric field polarized along the x axis. Equation (3) remains the same, Eq. (4) is zero, and Eq. (5) becomes

$$E_z(x,y,z) = \frac{1}{2\pi} \int \int_A E_x^{inc}(x',y',0) \frac{exp(ikR)}{R} \left(ik - \frac{1}{R} \right)\left(\frac{x-x'}{R} \right) dx'dy' \qquad (6)$$

In common situations we can neglect E_z in comparison with E_x. Hence, we assert the following statement: Eq. (3) represents the diffracted field in the right half-space from an aperture in a perfectly conducting plane screen. Determination of this cartesian component entails an evaluation of an integral expression. Additionally, we note that Eq. (3) is an exact solution to a specific boundary value problem.

$$(\nabla^2 + k^2)E_x(x,y,z) = 0 \quad , \qquad (7)$$

$$E_x(x,y,0) = F(x,y) \equiv \begin{cases} E_x^{inc}(x,y,0) & (x,y) \in A \\ \\ 0 & otherwise \end{cases} \qquad (8)$$

Additionally, we require that E_x satisfy the Sommerfeld radiation condition:

$$\lim_{R\to\infty} R\left(\frac{\partial E_x}{\partial z} - ikE_x \right) = 0 \quad . \qquad (9)$$

In other words E_x must vanish at least as fast as a diverging spherical wave for large R. That E_x satisfies Eqs. (7)–(9) can be verified by direct substitution, but it is not hard to understand in light of the origins of the derivation of Eq. (3). From Maxwell's equations for a homogeneous medium, each component of the electric field must satisfy the free-space scalar Helmholtz equation. The assumption that the value of the incident field in the aperture approximates sufficiently accurately the true aperture electric field establishes the boundary value described in Eq. (8).

Some simplification increases greatly the utility of Eq. (3). The factors $1/R$ in Eq. (3) will not differ significantly from $1/z$ under the assumption that the distance z between the aperture and the observation plane is much larger than the maximum dimension of the aperture and the observation region. (Hence, the obliquity factor z/R is nearly unity.) In practice $ik \gg 1/R$ so that

$$\left(ik - \frac{1}{R} \right) \approx ik \quad . \tag{10}$$

A further simplification suggested by Fresnel involves a binomial expansion of the factor R in the exponent of Eq. (3). This factor must be handled with greater care than the previous ones because of the highly oscillatory nature of $exp(ikR)$.

$$R = \left| (x-x')^2 + (y-y')^2 + z^2 \right|^{1/2} \quad ,$$

$$R \approx z \left| 1 + \frac{(x-x')^2 + (y-y')^2}{2z^2} \right| \quad . \tag{11}$$

Although it is an overly restrictive assumption, we can be sure this approximation is valid if

$$z^3 \gg \frac{\pi}{4\lambda} \left| (x-x')^2 + (y-y')^2 \right|^2_{MAX} \quad . \tag{12}$$

Substitution of the preceding simplifications and assumptions into Eq. (3) yields

$$E_x(x,y,z) = \frac{exp(ikz)}{i\lambda z} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(x',y') exp\left\{ \frac{i\pi}{\lambda z} \left| (x-x')^2 + (y-y')^2 \right| \right\} dx'dy' \quad . \tag{13}$$

The limits of integration are $\pm\infty$ because the aperture is accounted for by $F(x',y')$. Equation (13) is the Fresnel diffraction formula for diffraction of a

normally incident, linear x polarized wave by an aperture in an infinitely thin, perfectly conducting plane screen.

[Using the angular spectrum representation, we can obtain the Fresnel diffraction formula also. In his published notes Dainty (1983, §3.13-3.15) presents the pertinent steps. The angular spectrum representation follows from Eq. (3) using the Weyl formula or directly from the boundary value problem Eqs. (7)–(9). Restricting the angular spectrum to low frequency components leads to Eq. (13).]

As a result of the approximations employed in deriving Eq. (13), $E_x(x,y,z)$ no longer satisfies Eq. (7) exactly; a different differential equation is associated with Eq. (13). The first step in discovering this differential equation is to divide out the rapidly varying complex exponential factor, $exp(ikz)$, yielding a modified Fresnel integral.

$$e_x(x,y,z) = E_x(x,y,z)/exp(ikz) \quad , \tag{14}$$

$$e_x(x,y,z) = \frac{1}{i\lambda z} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(x',y') exp\left\{ \frac{i\pi}{\lambda z} \left[(x-x')^2 + (y-y')^2 \right] \right\} dx'dy' \quad . \tag{15}$$

Now by direct substitution one can show that Eq. (15) is a solution to the following parabolic equation (Feiock, 1978):

$$(\nabla^2_{xy} + 2ik \frac{\partial}{\partial z}) e_x(x,y,z) = 0 \quad , \tag{16}$$

subject to the boundary condition

$$e_x(x,y,0) = F(x,y) \tag{17}$$

Equation (16) is a slowly-varying envelope approximation to the Helmholtz equation. In other words, Eq. (16) is a simplification of Eq. (7) that is obtained by substituting Eq. (14) into Eq. (7) and assuming $e_x(x,y,z)$

55

is a slowly varying function of z compared with $exp(ikz)$ (Kogelnik and Li, 1965). Specifically, we assume

$$| \frac{\partial^2}{\partial z^2} e_x(x,y,z)| \ll k\,|\frac{\partial}{\partial z} e_x(x,y,z)| \quad , \tag{18}$$

i.e., the change in $\partial e_x/\partial z$ in a distance λ is small compared to $|\partial e_x/\partial z|$ itself. Evaluating Eq. (15) in the limit $z \rightarrow 0$ yields the boundary condition, Eq. (17). This can be shown in a straightforward manner if the following substitutions are made:

$$x' - x = p\,(i\lambda z/n)^{1/2} \quad ,$$
$$y' - y = s\,(i\lambda z/n)^{1/2} \quad . \tag{19}$$

Before continuing, we note that Eq. (16) has the same structure as the two-dimensional heat conduction equation (HCE) and the time-dependent Schrödinger equation. The one-dimensional HCE subject to an initial condition is solved by means of the Poisson-Weierstrass integral in the book by Pogorzelski (1966). The solution for the two-dimensional HCE is analogous. Likewise, the modified Fresnel diffraction integral, Eq. (15), has the same structure as the two-dimensional Poisson-Weierstrass integral.

Equation (14) and the succeeding assumption, Eq. (18), constitute the familiar definition of a beam field. In this definition we state the E_x is similar to a plane wave; however, the amplitude distribution is not uniform, but is concentrated near the axis of propagation, and the phase front is curved slightly. Equation (14) shows that the function e_x is the ratio of a beam and a plane wave.

Typically, in performing diffraction calculations, we are interested in evaluating Eq. (13) or equivalently Eq. (15). The relation of Eq. (15) to the

boundary value problem described by Eqs. (16) and (17) suggests an alternate method for obtaining a solution for the diffraction problem: expansion of the solution in the eigenfunctions of the differential equation with the expansion coefficients determined by the boundary condition.

The eigenfunctions of Eq. (16) have been studied for many years; they are the well known Gaussian beam modes. For systems with rectangular (x,y,z) geometry, the appropriate functions are Hermite-Gaussian (HG) functions, and for systems with cylindrical (r,θ,z) geometry, the appropriate functions are Laguerre-Gaussian (LG) functions. Because the corner of arbitrary angle considered in this chapter is described conveniently in a cylindrical geometry, the ensuing discussion utilizes LG functions. The derivation for a rectangular geometry follows in a similar fashion.

The notation for an LG function at $z = 0$, not necessarily the beam waist, is

$$\phi_{pl}(r',\theta',0) = \left(\frac{2r'^2}{w'^2} \right)^{|l|/2} L_p^{|l|} \left| \frac{2r'^2}{w'^2} \right| exp \left| r'^2 \left(-\frac{1}{w'^2} + \frac{in}{\lambda R'} \right) + il\theta' \right| \qquad , \qquad (20)$$

$$p = 0,1,2,\cdots,\infty \quad ; \quad l = -\infty,\cdots,-2,-1,0,1,2,\cdots,\infty \quad ,$$

where $2w'$ is the 1/e width or beam diameter of the Gaussian profile and R is the radius of curvature of the wavefront. L_p^l is an associated Laguerre polynomial. The eigenfunction is commonly designated TEM_{pl} where p and l are the radial and angular quantum numbers, respectively. (The presence of absolute value symbols around the angular number, $|l|$, is necessary because l can be negative. Other authors choose to denote the LG functions as a Fourier series in $cos(l\theta')$ and $sin(l\theta')$, and thus, they restrict l such that $l \geq 0$. Of course, a Fourier series in $exp(il\theta')$ is an

equivalent representation.) The observation plane is perpendicular to the z axis; the coordinates (r', θ') specify a point in this transverse plane.

Two properties of these functions will be important. First, these eigenfunctions form a complete and orthogonal set of functions. At least formally, any arbitrary distribution can be expanded in terms of these functions. Second, the functional form of each eigenfunction is the same for any value of z. In other words, an LG function retains its LG shape after propagation through some distance z.

These two properties are the foundation of and motivation for the Gaussian beam expansion approach. The first property implies the following expression

$$F(x',y') = F(r',\theta') = \sum_{p=0}^{\infty} \sum_{l=-\infty}^{\infty} C_{pl} \phi_{pl}(r',\theta',0) \quad . \tag{21}$$

(Although not mathematically precise, the notation in Eq. (21) is convenient. When a function depends on (x',y'), those variables will be specified; when the dependence is more logically on (r',θ'), those variables will be specified.) The C_{pl} are defined by

$$C_{pl} = \frac{\int_{-\pi}^{\pi} \int_0^{\infty} F(r',\theta') \quad \phi_{pl}^*(r',\theta',0) \; r'dr'd\theta'}{\int_{-\pi}^{\pi} \int_0^{\infty} \phi_{pl}(r',\theta',0) \phi_{pl}^*(r',\theta',0) \; r'dr'd\theta'} \quad . \tag{22}$$

Using Eq. (15) written in cylindrical coordinates and Eq. (21), $e_x(r,\theta,z)$ can be expressed as

$$e_x(r,\theta,z) = \frac{1}{i\lambda z} \int_{-\pi}^{\pi} \int_0^{\infty} \sum_{p=0}^{\infty} \sum_{l=-\infty}^{\infty} C_{pl} \phi_{pl}(r',\theta',0)$$

$$\times exp\left\{ \frac{i\pi}{\lambda z} \left[r'^2 - 2rr'\cos(\theta - \theta') + r^2 \right] \right\} r'dr'd\theta' \tag{23}$$

Reversing the order of the summation and the integration yields

$$e_x(r,\theta,z) = \sum_{p=0}^{\infty} \sum_{l=-\infty}^{\infty} C_{pl}\phi_{pl}(r,\theta,z) \quad , \tag{24}$$

where

$$\phi_{pl}(r,\theta,z) = \frac{1}{i\lambda z}\int_{-\pi}^{\pi}\int_{0}^{\infty}\phi_{pl}(r',\theta',0)\,exp\left\{\frac{i\pi}{\lambda z}\left|r'^{\,2}-2rr'\cos(\theta-\theta')+r^2\right|\right\}r'dr'd\theta' \quad . \tag{25}$$

Equation (25) is encountered when computing the Fresnel propagation of an LG function through a distance z. Thus, the second property is useful: $\phi_{pl}(r,\theta,z)$ has the same functional form as $\phi_{pl}(r,\theta,0)$. The beam radius, w', and the wavefront curvature, R', change as a function of z and a complex factor needs to be included. Straightforward evaluation of Eq. (25) yields

$$\phi_{pl}(r,\theta,z) = \frac{w'}{w}\left(\frac{2r^2}{w^2}\right)^{|l|/2}L_p^{|l|}\left|\frac{2r^2}{w^2}\right|$$

$$\times exp\left|r^2\left(-\frac{1}{w^2}+\frac{i\pi}{\lambda R}\right)+il\theta-i\delta(2p+|l|+1)\right| \quad , \tag{26}$$

where

$$w^2 = w'^{\,2}\left|\left(\frac{\lambda z}{\pi w'^{\,2}}\right)^2+\left(1+\frac{z}{R'}\right)^2\right| \quad , \tag{27}$$

$$R = z\left|\left(\frac{\pi w'^{\,2}}{\lambda z}\right)^2\left(1+\frac{z}{R'}\right)^2+1\right| / \left|1+\frac{z}{R'}\left(\frac{\pi w'^{\,2}}{\lambda z}\right)^2\left(1+\frac{z}{R'}\right)\right| \quad , \tag{28}$$

$$tan(\delta) = 1 / \left|\left(\frac{\pi w'^{\,2}}{\lambda z}\right)\left(1+\frac{z}{R'}\right)\right| \quad . \tag{29}$$

Note that the variation of w, R, and δ as a function of z is the same for all of the eigenfunctions, but that the phase shift, $-i\delta(2p+l+1)$, is dependent on the quantum number. Equation (24) with coefficients (22) and LG functions (26) constitute the Gaussian beam expansion approach.

The validity of term-by-term integration of Eq. (23) is an important point to consider. We must investigate the convergence of the series, Eq. (21). Uniform convergence of Eq. (21) is a sufficient condition to permit the step from Eq. (23) to Eq. (24). If the function $F(r',\theta')$ is a continuous function on $r' = (0,\infty)$ and $\theta' = (-\pi,\pi)$ with a piecewise continuous first derivative on that interval, then, indeed, the convergence is absolute and uniform. If, however, $F(r',\theta')$ is not continuous, then the series in Eq. (23) cannot converge uniformly because a uniformly convergent series of continuous functions, e.g., ϕ_{pl}, always converges to a continuous function. Nevertheless, uniform convergence is only sufficient, not necessary. There are relaxed conditions under which term-by-term integration is permissible. For example, switching the integration and the summation is valid if the series is uniformly convergent almost everywhere, i.e., except on a set of measure zero, and is boundedly convergent over the whole interval (Titchmarsh, 1932). In any event, because the approximations employed in deriving the paraxial scalar wave equation exclude very high order modes, the summations cannot be extended over infinite limits. In other words, at some very high mode number, the previously stated conditions are violated. Also, practical use of this Gaussian beam expansion for computer calculations will require the summations to be finite. In that case, the reversal of the sum and the integral is valid.

To denote the finite limits of the summation in Eq. (24), we rewrite it so that $0 \le p \le P$ and $-L \le l \le L$:

$$e_x(r,\theta,z) = \sum_{p=0}^{P} \sum_{l=-L}^{L} C_{pl} \phi_{pl}(r,\theta,z) \quad . \tag{30}$$

3.1.2   Calculation of the Expansion Coefficients
and Computational Results

3.1.2.1   Uniform Circular Aperture

Before considering the diffraction from a corner of arbitrary angle problem, it is helpful to understand some of the operating characteristics of the Gaussian beam expansion. Analysis of a simpler diffraction problem will provide some of these insights. Hence, we consider diffraction by a circular aperture first.

Assume that a converging spherical wave of radius $z_0 > 0$ illuminates a circular aperture of radius a. The incident field, written in the paraxial approximation and scaled by z_0, is

$$E_x^{inc}(r,\theta,z) = z_0 \frac{exp[i\pi r^2/\lambda(z-z_0)]}{|z-z_0|} exp[ik(z-z_0)] = e_x^{inc}(r,\theta,z)exp[ik(z-z_0)] \quad , \quad (31)$$

where z_0 is the radius of curvature of the phase front at $z = 0$. The GBE does not require that the illumination be a converging beam ($z_0 > 0$); the assumption introduced here is illustrative. The appropriate boundary value function is

$$F(r',\theta') = circ\left(\frac{r'}{a}\right)e_x^{inc}(r',\theta',0) = circ\left(\frac{r'}{a}\right)exp[-i\pi r'^2/\lambda z_0] \quad , \quad (32)$$

where

$$circ\left(\frac{r'}{a}\right) = \begin{cases} 1 & 0 \leq r' \leq a \\ 0 & r' > a \end{cases} \quad -\infty < r' < \infty \quad .$$

We evaluate the expansion coefficients C_{pl} in a straightforward manner. Inserting Eqs. (20) and (32) into Eq. (22) gives

$$C_{pl} = \frac{\int_{-\pi}^{\pi}\int_0^a \left(\frac{2r'^2}{w'^2}\right)^{|l|/2} L_p^{|l|}\left[\frac{2r'^2}{w'^2}\right] exp\left[-\frac{r'^2}{w'^2} - il\theta'\right] r'dr'd\theta'}{N_{pl}} \quad , \quad (33)$$

where

$$N_{pl} \equiv \int_{-\pi}^{\pi} \int_{0}^{\infty} |\phi_{pl}(r',\theta',0)|^2 \, r' dr' d\theta' = 2\pi \frac{w'^2}{4} \frac{(|l|+p)!}{p!} \quad , \tag{34}$$

and we have chosen the value $R' = -z_0$. It is logical to choose the phase front curvatures to be the same. Since the boundary value function does not depend on θ', only coefficients for $l = 0$ are nonzero. In other words,

$$C_{pl} = \delta_{0l} \frac{4}{w'^2} \int_0^a L_p^0 \left| \frac{2r'^2}{w'^2} \right| exp \left| -\frac{r'^2}{w'^2} \right| r' dr' \quad . \tag{35}$$

δ_{0l} is a Kronecker delta function, which equals 1 when $l = 0$ and equals 0 when $l \neq 0$. A simple variable substitution, $\xi \equiv 2r'^2/w'^2$, in Eq. (35) yields

$$C_{pl} = \delta_{0l} \int_0^{2a^2/w^2} L_p^0(\xi) \, exp(-\xi/2) \, d\xi \quad . \tag{36}$$

We can evaluate Eq. (36) analytically by expanding $L_p^0(\xi)$ in a finite power series and integrating term-by-term, but numerical integration of Eq. (36) is straightforward, accurate, and relatively easy. If we consider the coefficients to have been evaluated, then the GBE solution for this diffraction problem is a single sum over the index p

$$e_x^{GBE}(r,\theta,z) = \sum_{p=0}^{P} C_{p0} \phi_{p0}(r,\theta,z) \quad , \tag{37}$$

with

$$\phi_{p0}(r,\theta,z) = \frac{w'}{w} L_p^0 \left| \frac{2r^2}{w^2} \right| exp \left| r^2 \left(-\frac{1}{w^2} + \frac{i\pi}{\lambda R} \right) - i\delta(2p+1) \right| \quad . \tag{38}$$

The parameters w, R, and δ are dependent on their values at $z = 0$ (w' and R') and the distance z [see Eqs. (27)–(29)].

Before comparing this result to the closed-form circular aperture solution obtained directly from the modified Fresnel diffraction integral Eq. (15), we investigate how well the GBE reproduces the boundary value function. At $z = 0$

$$e_x^{GBE}(r',\theta',0) = \left\{ \sum_{p=0}^{P} C_{p0} L_p^0 \left[\frac{2r'^2}{w'^2} \right] exp\left[-\frac{r'^2}{w'^2} \right] \right\} exp\left[-\frac{i\pi r'^2}{\lambda z_0} \right] \quad (39)$$

The complex factor $exp(-i\pi r'^2/\lambda z_0)$ is present in Eqs. (32) and (39); hence, we compare the square modulus of both equations.

The computational behavior of this expansion will depend on the number of terms retained, $P + 1$, and on the value chosen for the beam width w'. Determination of an optimal w' is a complex subject to address. Bogush and Elkins (1986) established one optimization criterion. The beam width or equivalently the scale can be optimized such that the radiated power of the approximated distribution best matches the radiated power of the actual distribution. This criterion is especially important when one is trying to use a very limited number of terms in the expansion, say, less than ten. Since choosing an optimal w' will not be greatly relevant to the calculations for the corner diffraction problem, we do not discuss this topic here. For the problem at hand, we choose

$$w' = (2)^{1/2} a \quad (40)$$

The dependence of the GBE on the number of terms $P + 1$ is compared in Fig. 3.3. Also shown is the true boundary value function. The square modulus of e_x is plotted versus the radial coordinate normalized to the radius of the aperture, i.e., r'/a. Increasing the number of terms from 20 to 50 improves the approximation somewhat, i.e., the amplitude of the oscillations about the true boundary value function decreases. In contrast, however, increasing the number of terms to 100 does not decrease the amplitude of these oscillations although they increase in number and the drop-off at the edge of the aperture is steeper. The convergence behavior is reminiscent of a Fourier series approximation to a square pulse function:
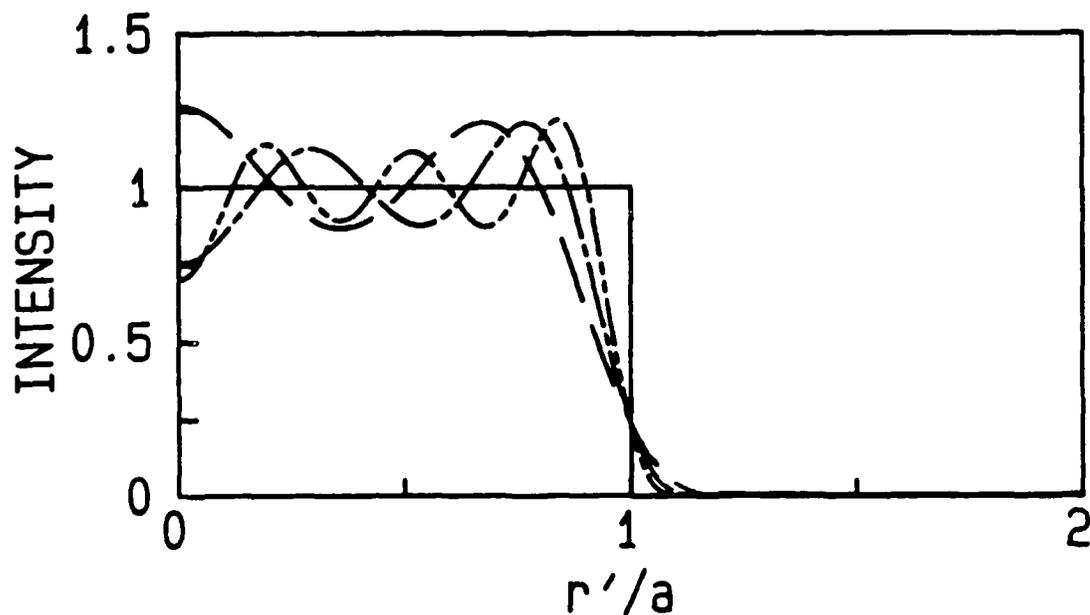
Fig. 3.3. GBE representation of uniform circular aperture, *circ(r'/a)*. The
solid line (———) represents the intensity of the input function plotted vs a
normalized radial coordinate. The dashed curves depict the GBE
approximation for P = 20 (- - - -), P = 50 (- - – -), and P = 100 (– - - – - -).

the GBE value of the field at the edge is 0.5 (the intensity is 0.25), and the

oscillations do not decrease at the discontinuous step. On the other hand

the Gibb's phenomenon at the edge is less pronounced for the GBE than for

a Fourier series because of the Gaussian dependence of the eigenfunctions.

Although we might question the quality of the representation of the

aperture distribution by the GBE, we should bear in mind that the desire

here is to approximate accurately the diffraction pattern. Obviously, the

discontinuity in the boundary value function in the aperture plane makes

it difficult for the expansion to converge quickly to the precise distribution,

but the diffraction pattern has a smoother behavior. That the GBE closely

approximates the diffraction pattern will be apparent in the following discussion.

The focal plane of the incident converging wavefront, $z = z_0$, also known as the Fourier plane, is an important place to compare the diffraction patterns (see Fig. 3.2). The field in this plane represents, essentially, the Fourier transform of the aperture function. For the case of the uniformly illuminated circular aperture, we obtain the field by substituting the boundary value function, Eq. (32), into the modified Fresnel diffraction integral Eq. (15) and computing the analytic result. The result is the well known Airy disk pattern:

$$e_x^{AIRY}(r,\theta,z_0) = \left(\frac{na^2}{i\lambda z_0} \right) exp\left| \frac{inr^2}{\lambda z_0} \right| \left| 2\frac{J_1(2nar/\lambda z_0)}{(2nar/\lambda z_0)} \right| \quad . \tag{41}$$

Before comparing the precise result to the GBE, we evaluate Eq. (37) at $z = z_0$. Primarily, we have to compute w, R, and δ by evaluating Eqs. (27)–(29) at the distance $z = z_0$. ($R' = -z_0$ from before.) At this position the results are easy to obtain:

$$w = \frac{\lambda z_0}{nw'} \quad ,$$

$$R = z_0 \quad , \tag{42}$$

$$\delta = n/2 \quad .$$

Noting that the choice of w' in Eq. (40) gives

$$\frac{w'}{w} = \frac{nw'^2}{\lambda z_0} = 2\frac{na^2}{\lambda z_0} \quad ,$$

and defining a normalized radial coordinate

$$\chi \equiv \frac{(2)^{1/2} r}{w} = \frac{kar}{z_0} \quad , \tag{43}$$

where $k = 2\pi/\lambda$, Eqs. (37) and (38) become

$$e_x^{GBE}(\chi,\theta,z_0) = \left(\frac{\pi a^2}{i\lambda z_0}\right) exp\left|\frac{i\lambda z_0 \chi^2}{4\pi a^2}\right| \left\{2\sum_{p=0}^{P} C_{p0}L_p^0(\chi^2)exp(-\chi^2/2)(-1)^p\right\} \quad . \quad (44)$$

Rewriting Eq. (41) in terms of the normalized radial coordinate yields

$$e_x^{AIRY}(\chi,\theta,z_0) = \left(\frac{\pi a^2}{i\lambda z_0}\right) exp\left|\frac{i\lambda z_0 \chi^2}{4\pi a^2}\right| \left\{2\frac{J_1(\chi)}{\chi}\right\} \quad . \quad (45)$$

The square modulus of the bracketed factors {} in Eqs. (44) and (45) is proportional to the observed optical intensity in the Fourier plane. We will examine the optical intensity of the diffraction pattern normalized to unity at the center of the pattern. Figure 3.4 compares the normalized intensity of Eq. (44) for P = 20, 50, and 100 to the normalized intensity in Eq. (45).
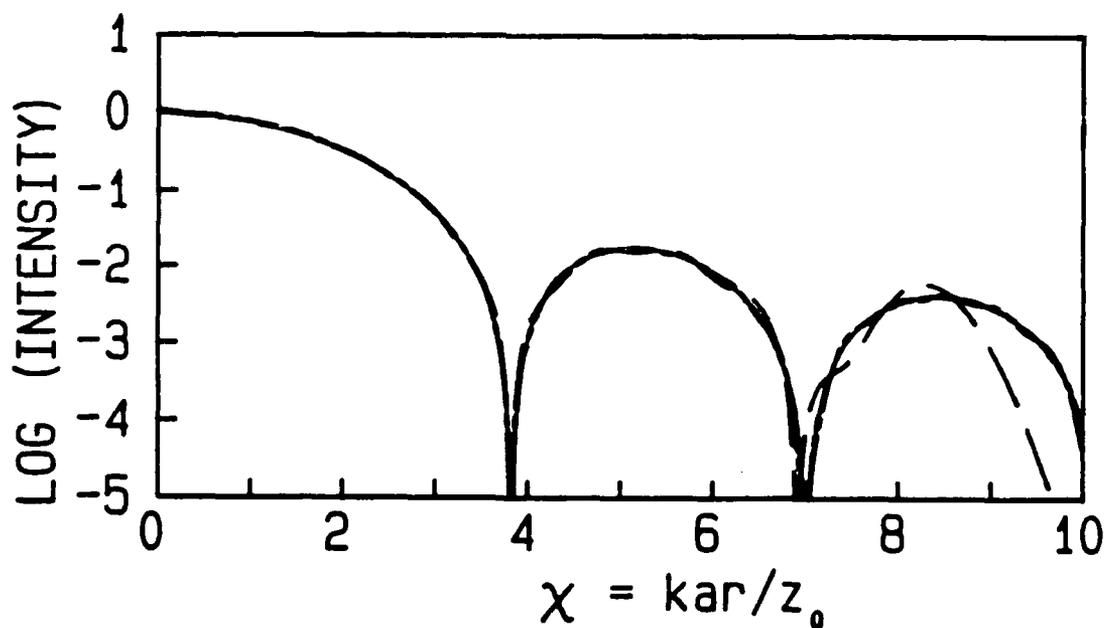

Fig. 3.4. GBE representation for diffraction pattern of uniform circular aperture. The logarithm of intensity is plotted vs a normalized radial coordinate. The solid line (———) represents the precise solution, i.e., the Airy disk function. The dashed curves depict the GBE approximation for P = 20 (- - - -), P = 50 (— - — -), and P = 100 (— - - — - -).

The vertical axis is log(intensity) in order to bring out the low intensity differences between the curves. With P = 20, the GBE is virtually indistinguishable from the precise calculation over the central bright spot, but the deviations are discernible in the first ring. The second zero of the pattern is visibly wrong. Increasing to P = 50 improves the representation; deviations are apparent in the second bright ring. P = 100 is an accurate representation over the entire range shown.

As we stated earlier, the convergence of the GBE to the precise analytic result for the diffraction pattern is much more acceptable than the convergence to the aperture distribution. The following theorem provides some insight into this differing convergence behavior (Goldberg, 1965, page 6, Theorem 3B).

"If $f, f_1, f_2, \cdots$ are in L^1 and if $\|f_n - f\|_1 \to 0$ as $n \to \infty$, then

$$\lim_{n \to \infty} \bar{f}_n(x) = \bar{f}(x) \quad uniformly \ (-\infty < x < \infty) \quad .$$

[The function f on $(-\infty, \infty)$ is said to be of class L^1 ($f \in L^1$) if

$$\int_{-\infty}^{\infty} |f(x)| \, dx < \infty \quad .$$

If $f \in L^1$, then

$$\|f\|_1 \equiv \int_{-\infty}^{\infty} |f(x)| \, dx \quad ,$$

i.e., the L^1 norm of f. The Fourier transform $\bar{f}$ of $f \in L^1$ is defined by

$$\bar{f}(x) = \int_{-\infty}^{\infty} f(t) \exp(ixt) \, dt \quad ,$$

for $f \in L^1$ and $(-\infty < x < \infty)$.]"

In the case of the finite circular aperture, and indeed for the corner
aperture considered next, the boundary value function F is certainly a
member of L^1. The LG functions are L^1 as well. That the GBE converges in
the mean of order 1 is not as obvious but follows from another theorem
(Goldberg, 1965, page 4, Theorem 2H).

"Let $f, f_1, f_2, \cdots$ be in L^1. If

$$\lim_{n \to \infty} f_n(x) = f(x) \quad almost\ everywhere\ (-\infty < x < \infty)$$

and

$$\lim_{n \to \infty} \|f_n\|_1 = \|f\|_1 \quad,$$

then

$$\lim_{n \to \infty} \|f_n - f\|_1 = 0 \quad."$$

The GBE is pointwise convergent almost everywhere; the only points of
nonconvergence are the discontinuities. The second condition of the
Theorem 2H is satisfied as well. Hence, the convergence of the Fourier
transform of the GBE for the aperture distribution, i.e., the GBE evaluated
at the Fourier plane, is uniform. Figure 3.4 is consistent with this kind of
convergence.

3.1.2.2 Gaussian Illuminated Corner

In the sample calculation just discussed we discovered some of the
basic characteristics of the GBE. In this chapter, however, we desire to
solve a new problem: diffraction of an incident light wave by a corner of
arbitrary angle. The specific problem geometry was illustrated in Fig. 3.1.
The corner has an opening angle θ_0 cut in a thin, perfectly conducting
screen. Allow an LG beam ϕ_{qm} to illuminate the corner. In particular the

boundary value function is

$$F(r',\theta') = rect\left(\frac{\theta'}{\theta_0}\right)\phi_{qm}(r',\theta',0) \quad , \tag{46}$$

where

$$rect\left(\frac{\theta'}{\theta_0}\right) = \begin{cases} 1 & |\theta'| \le \theta_0/2 \\ 0 & |\theta'| > \theta_0/2 \end{cases} \quad -n < \theta' < n \tag{47}$$

and $\phi_{qm}(r',\theta',0)$ is defined in Eq.(20).

As was seen in the circular aperture problem, computation of the expansion coefficients solves the diffraction problem. In this problem the expansion coefficients are expressible in closed-form. Proceeding as before,

$$C_{pl} = \frac{1}{N_{pl}} \int_{-n}^{n} \int_{0}^{\infty} \left\{ rect\left(\frac{\theta'}{\theta_0}\right)\left(\frac{2r'^2}{w'^2}\right)^{|m|/2} L_q^{|m|}\left[\frac{2r'^2}{w'^2}\right] exp\left[-\frac{r'^2}{w'^2} + im\theta'\right] \right\}$$

$$\times \left\{ \left(\frac{2r'^2}{w'^2}\right)^{|l|/2} L_p^{|l|}\left[\frac{2r'^2}{w'^2}\right] exp\left[-\frac{r'^2}{w'^2} - il\theta'\right] \right\} r'dr'd\theta' \tag{48}$$

For this diffraction problem we set equal the beam widths and phase fronts, i.e., $w'_{GBE} = w'_{inc} = w'$ and $R'_{GBE} = R'_{inc} = R'$.

We can evaluate the θ' integration in Eq. (48) easily.

$$g_l(\theta_0) = \int_{-\theta_0/2}^{\theta_0/2} exp[i(m-l)\theta']d\theta' = \theta_0 \frac{sin[(m-l)\theta_0/2]}{(m-l)\theta_0/2} \quad . \tag{49}$$

The r' integration is somewhat more difficult. First, write Eq. (48) as

$$C_{pl} = \frac{1}{N_{pl}} g_l(\theta_0) K_{qmpl} \quad , \tag{50}$$

where

$$K_{qmpl} = \int_{0}^{\infty} \left\{ \left(\frac{2r'^2}{w'^2}\right)^{|m|/2} L_q^{|m|}\left[\frac{2r'^2}{w'^2}\right] exp\left[-\frac{r'^2}{w'^2}\right] \right\}$$

$$\times \left\{ \left(\frac{2r'^2}{w'^2} \right)^{|l|/2} L_p^{|l|} \left| \frac{2r'^2}{w'^2} \right| exp \left| -\frac{r'^2}{w'^2} \right| \right\} r'dr' \quad . \tag{51}$$

Make the substitution

$$\xi \equiv 2r'^2 / w'^2 \tag{52}$$

in Eq. (51) to give

$$K_{qmpl} = \frac{w'^2}{4} \int_0^\infty \xi^{|m|/2} L_q^{|m|}(\xi) \, \xi^{|l|/2} L_p^{|l|}(\xi) \, exp(-\xi) \, d\xi \quad . \tag{53}$$

We expand the associated Laguerre polynomial, $L_q^m(\xi)$, as a finite series in ξ (Gradshteyn and Ryzhik, 1980, 8.970.1), i.e.,

$$L_q^{|m|}(\xi) = \sum_{q'=0}^{q} (-1)^{q'} \binom{q+|m|}{q-q'} \frac{\xi^{q'}}{q'!} \quad . \tag{54}$$

Equation (53) becomes

$$K_{qmpl} = \frac{w'^2}{4} \sum_{q'=0}^{q} (-1)^{q'} \binom{q+|m|}{q-q'} \frac{1}{q'!} \int_0^\infty \xi^{|m|/2+|l|/2+q'} L_p^{|l|}(\xi) \, exp(-\xi) \, d\xi \quad . \tag{55}$$

The integral can be evaluated (Gradshteyn and Ryzhik, 1980, 7.414.11).

$$K_{qmpl} = \frac{w'^2}{4} \sum_{q'=0}^{q} (-1)^{q'} \binom{q+|m|}{q-q'} \frac{1}{q'!} \frac{\Gamma\left(\frac{|m|}{2} + \frac{|l|}{2} + q' + 1 \right) \Gamma\left(p + \frac{|l|}{2} - \frac{|m|}{2} - q' \right)}{p! \, \Gamma\left(\frac{|l|}{2} - \frac{|m|}{2} - q' \right)} \tag{56}$$

Putting all the pieces together, Eqs. (34), (49), and (56) into Eq. (50), we obtain a closed-form solution for the coefficients C_{pl}:

$$C_{pl} = \frac{\theta_0}{2\pi} \frac{sin[(m-l)\theta_0/2]}{(m-l)\theta_0/2} \frac{1}{(|l|+p)!} \sum_{q'=0}^{q} (-1)^{q'} \binom{q+|m|}{q-q'} \frac{1}{q'!}$$

$$\times \frac{\Gamma\left(\frac{|m|}{2} + \frac{|l|}{2} + q' + 1 \right) \Gamma\left(p + \frac{|l|}{2} - \frac{|m|}{2} - q' \right)}{\Gamma\left(\frac{|l|}{2} - \frac{|m|}{2} - q' \right)} \quad . \tag{57}$$

Formally, computation of these coefficients completes the solution to the boundary value problem described by Eq. (46).

Now consider a simplification: the corner is illuminated by the elementary LG function, ϕ_{00}. Setting $q = m = 0$ in Eq. (57) yields

$$C_{p\ell} = \frac{\theta_0}{2\pi}\,\frac{sin[l\theta_0/2]}{l\theta_0/2}\,\frac{1}{(|l|+p)!}\,\frac{\Gamma\left(\frac{|l|}{2}+1\right)\Gamma\left(p+\frac{|l|}{2}\right)}{\Gamma\left(\frac{|l|}{2}\right)} \quad . \tag{58}$$

The recurrence formula for the Gamma function (Abramowitz and Stegun, 1972, 6.1.15) simplifies the equation.

$$C_{p\ell} = \left\{\frac{\theta_0}{2\pi}\,\frac{sin[l\theta_0/2]}{l\theta_0/2}\right\}\left\{\frac{1}{(|l|+p)!}\left(\frac{|l|}{2}\right)\Gamma\left(p+\frac{|l|}{2}\right)\right\} \quad . \tag{59}$$

If we write out the GBE solution at $z = 0$ in detail and rearrange factors intuitively, we find

$$e_x(r',\theta',0) = F(r',\theta') = rect\left(\frac{\theta'}{\theta_0}\right)exp(-r'^2/w'^2)\,exp(i\pi r'^2/\lambda R') \quad , \tag{60}$$

$$e_x(r',\theta',0) = \sum_{p=0}^{P}\sum_{l=-L}^{L}\left\{\frac{\theta_0}{2\pi}\,\frac{sin|l\theta_0/2|}{l\theta_0/2}\,exp(il\theta')\right\}$$

$$\times\left\{\frac{1}{(|l|+p)!}\left(\frac{|l|}{2}\right)\Gamma\left(p+\frac{|l|}{2}\right)\left(\frac{2r'^2}{w'^2}\right)^{|l|/2}L_p^{|l|}\left|\frac{2r'^2}{w'^2}\right|exp(-r'^2/w'^2)\right\}exp(i\pi r'^2/\lambda R') \quad . \tag{61}$$

The arrangement of factors in Eq. (61) makes clear how the LG functions combine to represent the boundary value function. The first bracketed factor describes the angular dependence of the boundary value function. In Eq. (60) replace $rect(\theta'/\theta_0)$ by its Fourier series expansion, i.e.,

$$rect\left(\frac{\theta'}{\theta_0}\right) = \sum_{l=-\infty}^{\infty}\frac{\theta_0}{2\pi}\,\frac{sin|l\theta_0/2|}{l\theta_0/2}\,exp(il\theta') \quad , \tag{62}$$

so that Eq. (60) becomes

$$F(r',\theta') = \sum_{l=-\infty}^{\infty}\frac{\theta_0}{2\pi}\,\frac{sin[l\theta_0/2]}{l\theta_0/2}\left(exp(il\theta')\,exp(-r'^2/w'^2)\right)exp(i\pi r'^2/\lambda R') \quad . \tag{63}$$

Now replace $exp(il\theta')exp(-r'^2/w'^2)$ by its GBE. Straightforward evaluation and use of integral tables (Gradshteyn and Ryzhik, 1980, 7.414.11) show that

$$exp(il\theta')\,exp(-r'^2/w'^2) = \sum_{p=0}^{\infty} \frac{1}{(|l|+p)!} \left(\frac{|l|}{2}\right)\Gamma\!\left(p+\frac{|l|}{2}\right)$$

$$\times \left(\frac{2r'^2}{w'^2}\right)^{|l|/2} L_p^{|l|}\left|\frac{2r'^2}{w'^2}\right| exp(-r'^2/w'^2)\,exp(il\theta') \quad . \tag{64}$$

Putting Eqs. (63) and (64) together gives Eq. (61). A summation over p of the second bracketed term in Eq. (61) is an expansion of $exp(-r'^2/w'^2)$ in terms of LG modes of a given index l. The separability of the r' and θ' dependence in the boundary value function is preserved in the GBE.

By virtue of the preceding discussion, we can understand the GBE solution at $z = 0$ by studying the two parts separately. Figure 3.5 shows the representation of $rect(\theta'/\theta_0)$ for $\theta_0 = \pi/2$ by its Fourier series expansion Eq. (62). The familiar ringing, especially near the points of discontinuity, is indicative of the mean-square convergence behavior of Fourier series. The representation of the square pulse distribution improves as the number of terms L is increased from 10 to 25 and finally to 100.

Figure 3.6 illustrates the behavior of Eq. (64) for $l = 1, 3$, and 5 and P = 25, 100, and 1000. [Appendix C: Calculations for Gaussian Beam Expansion discusses the accurate computation of the expansion coefficients in Eq. (64).] The amplitude of Eq. (64) is plotted vs a normalized radial argument, $\chi' = (2)^{1/2}r'/w'$ as in Eq. (43). For $l = 0$, there is only one term in the expansion: the elementary Gaussian mode $exp(-r'^2/w'^2)$. Hence, the representation is exact. For $l > 0$, the LG modes are zero at $\chi' = 0$;

therefore, it is difficult to approximate well the exact function, which
equals *one* near that point. As χ' increases, however, the GBE aproaches
the Gaussian profile. For larger values of P, i.e., more terms included in
the summation, the approximation is much improved for smaller values of
χ'.

We turn our attention now to the primary intent of this GBE
analysis: evaluation of diffraction patterns from corners. We carry out
this evaluation in the focal or Fourier plane. In other words, $R' = -z_0$ and z
$= z_0$ as for the uniform circular aperture problem. The LG function
parameters w, R, and δ are given by Eq. (42). The GBE gives
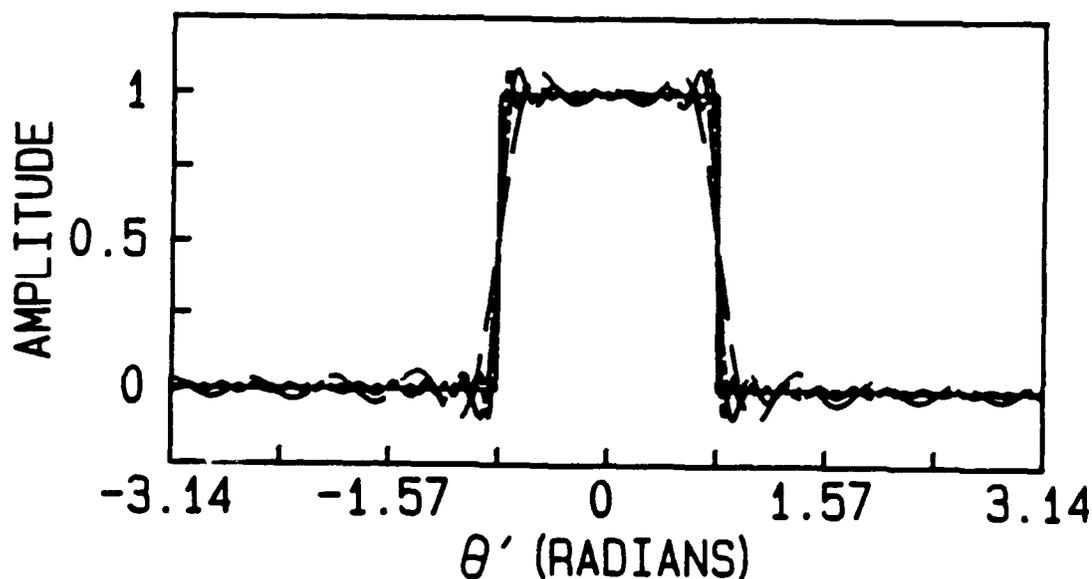
Fig. 3.5. GBE representation of plane-screen corner, *rect*(θ/θ_0). The
opening angle is $\theta_0 = \pi/2$, and the amplitude of the input function vs angle
is represented by the solid line (——). The dashed curves depict the GBE
approximation for L = 10 (- - - -), L = 25 (– – –), and L = 50 (– - – - –).
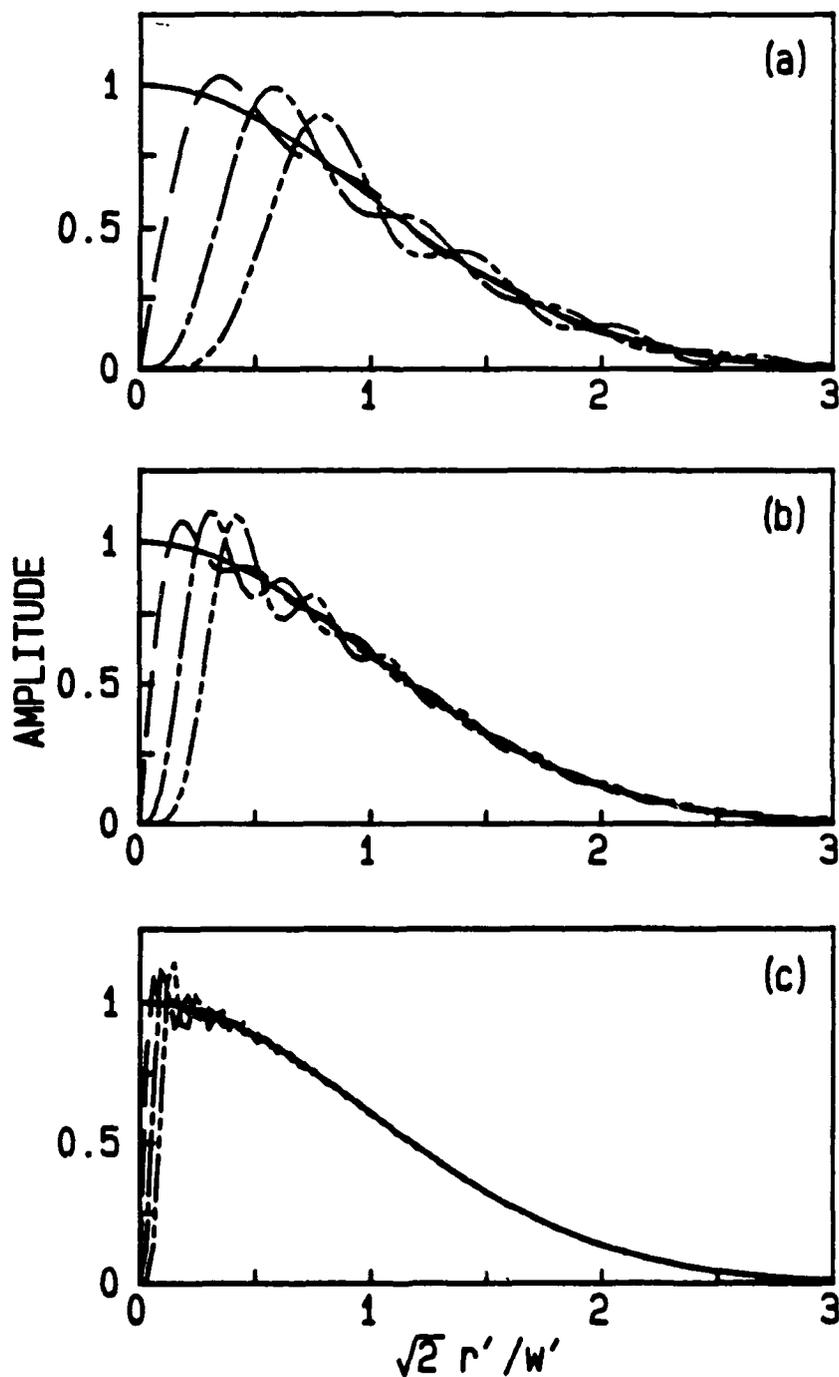
Fig. 3.6. GBE representation of radial dependence oi a Gaussian illuminated corner. The opening angle is $\theta_0 = \pi/2$, and the amplitude of the input function vs a normalized radial coordinate is represented by the solid line (——). The dashed curves depict the GBE approximation for $l = 1$ (- - - -), $l = 3$ (— - — -), and $l = 5$ (— - - — - -). The three sets of graphs were calculated for (a) P = 25, (b) P = 100, and (c) P = 1000.

$$e_x(\chi,\theta,z_0) = \frac{nw'^2}{i\lambda z_0} exp\left[\frac{i\lambda z_0 \chi^2}{2nw'^2}\right] \sum_{p=0}^{P} \sum_{l=-L}^{L} \left\{ \frac{\theta_0}{2n} \frac{sin[l\theta_0/2]}{l\theta_0/2} \frac{1}{(|l|+p)!} \left(\frac{|l|}{2}\right)\Gamma\left(p+\frac{|l|}{2}\right)\right\}$$

$$\times \chi^{|l|} L_p^{|l|}(\chi^2) exp(il\theta) exp(-\chi^2/2)(-i)^{2p+|l|} \quad , \quad (65)$$

where $\chi \equiv (2)^{1/2} r/w$ as in Eq. (43).

In Fig. 3.7 we present three-dimensional plots of some computational results from Eq. (65). The vertical direction is $log(|e_x(\chi,\theta,z_0)|^2/|e_x(0,\theta,z_0)|^2)$, i.e., the logarithm of the optical intensity normalized to unity at the center of the pattern. The values plotted are clipped at 40 dB down from this central peak intensity. The pattern is sampled on a cartesian grid; the cartesian coordinates are normalized in the same way as the radial coordinate in Eq. (43), that is,

$$(\chi_x,\chi_y) = ((2)^{1/2}(x/w),(2)^{1/2}(y/w)) \quad .$$

The values range from $\chi_{xy} \in (-10,10)$ on a 41×41 grid.

Shown are results for different opening angles: $n/3$ (a), $n/2$ (b), $3n/4$ (c), n (d), and $3n/2$ (e). With the exception of Fig. 3.7(d), which is the pattern from an edge, all of the patterns display two spikes or ridges of high intensity; Fig. 3.7(d) displays only one spike. These spikes are oriented at angles such that each spike is perpendicular to one of the edges in the aperture. In the $n/3$ corner pattern, for example, one spike runs along $\theta = 2n/3$ (and $-n/3$) and is perpendicular to the edge at $\theta = n/6$. The other spike runs along $\theta = -2n/3$ (and $n/3$) and is perpendicular to the edge at $\theta = -n/6$.

The figures were computed using the limits $P = 100$ and $L = 25$. These values were high enough to produce satisfactory results; more accurate results are possible, but the increase in computer time was
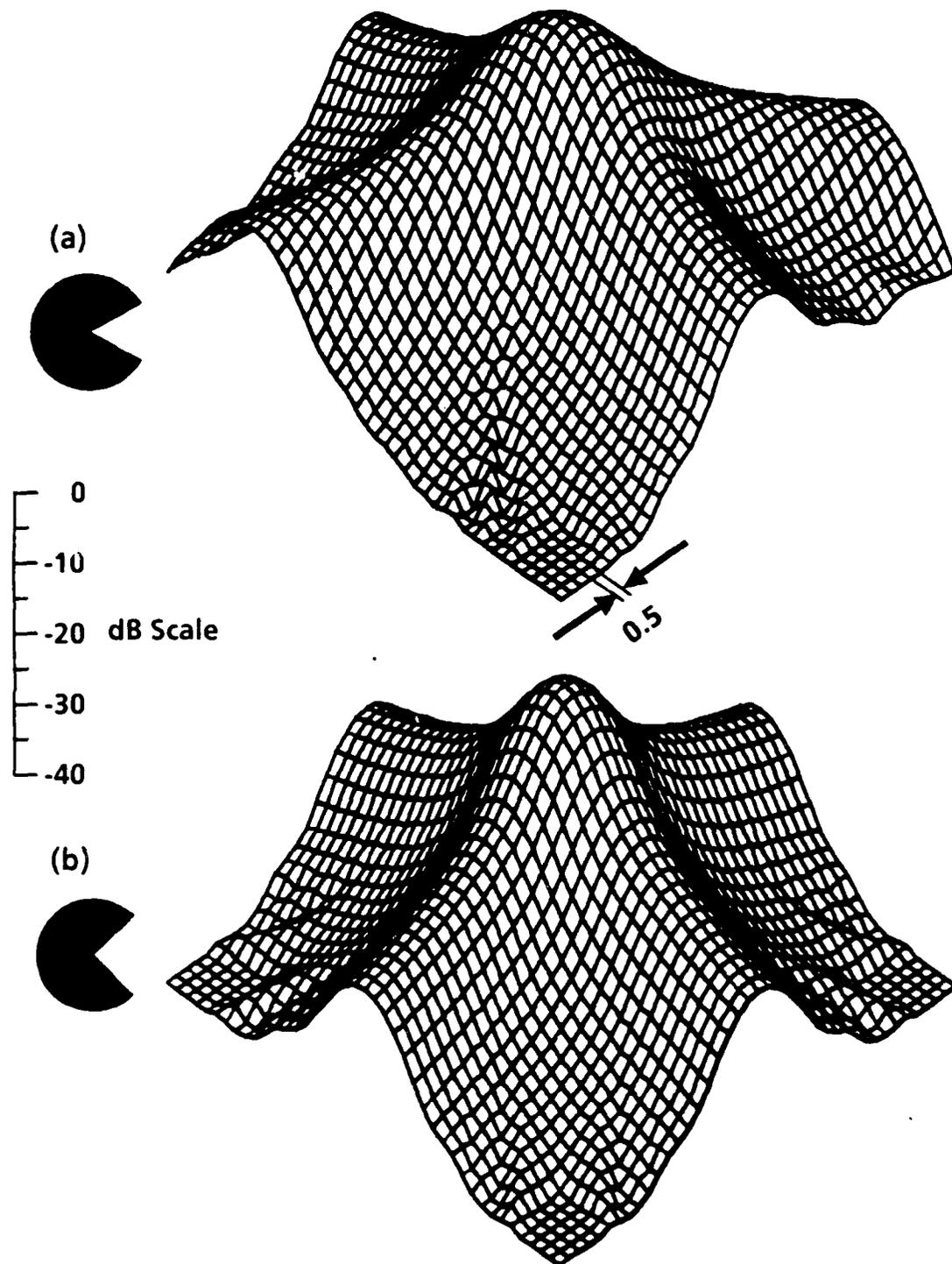
Fig. 3.7. Three-dimensional representations of the corner diffraction patterns. Shown are GBE calculations for the logarithm of the normalized optical intensity of the diffraction pattern produced by the following corners: $\theta_0 = \pi/3$ (a), $\pi/2$ (b). The corner apertures are shown at left.
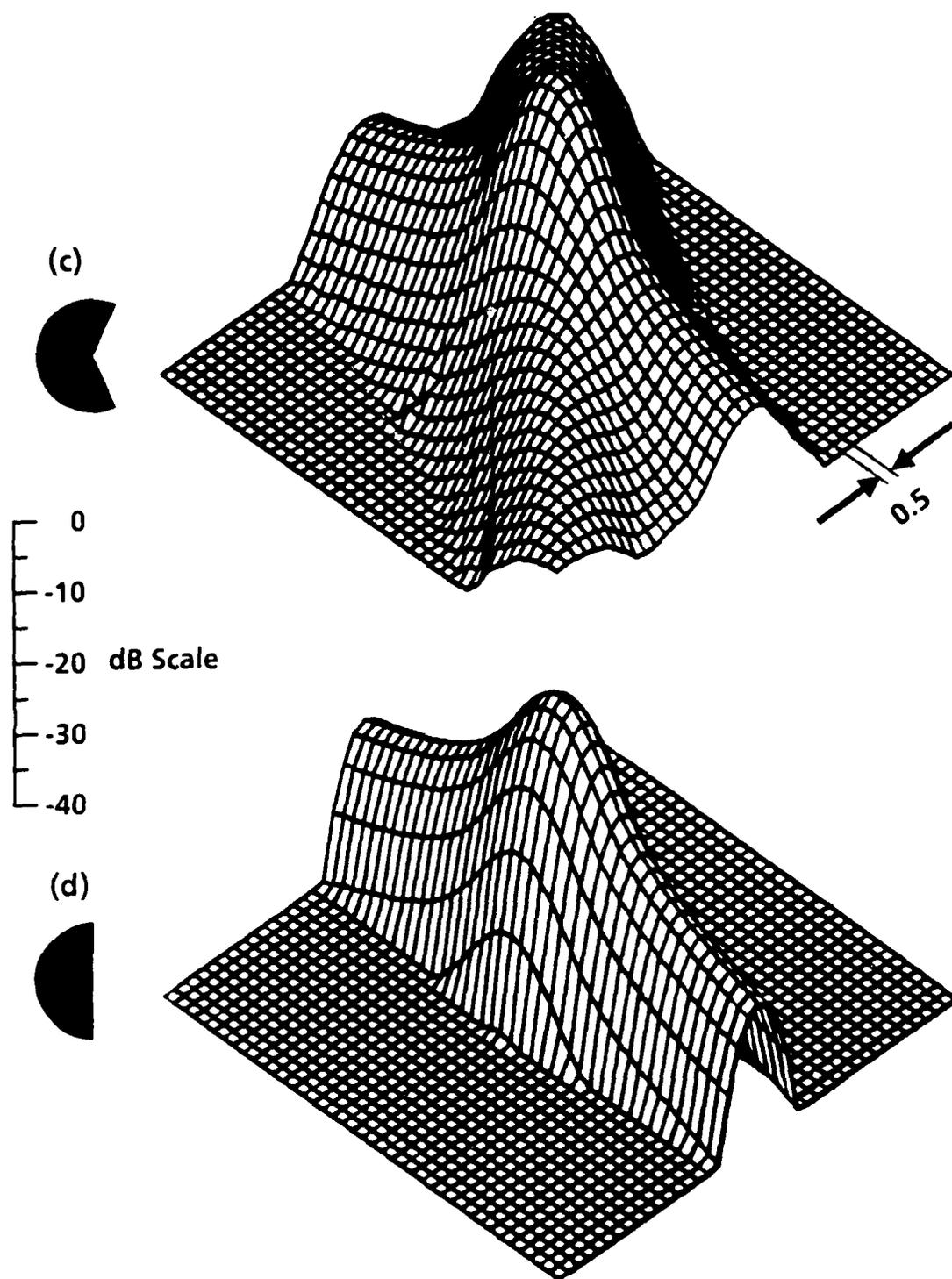
Fig. 3.7. Three-dimensional representations of the corner diffraction patterns. Shown are GBE calculations for the logarithm of the normalized optical intensity of the diffraction pattern produced by the following corners: $\theta_0 = 3\pi/4$ (c), π (d). The corner apertures are shown at left.

(e)

$$\begin{array}{l} 0 \\ -10 \\ -20 \quad \text{dB Scale} \\ -30 \\ -40 \end{array}$$

0.5
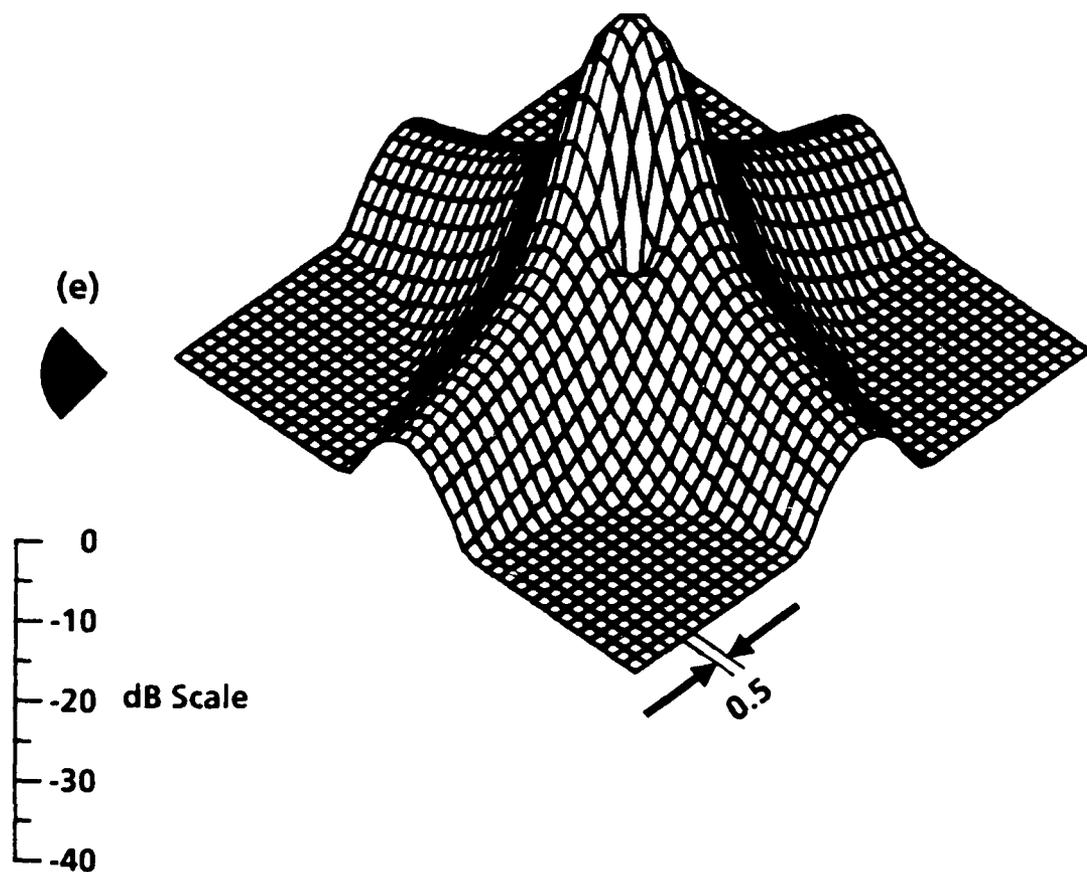
Fig. 3.7. Three-dimensional representation of the corner diffraction pattern. Shown is GBE calculation for the logarithm of the normalized optical intensity of the diffraction pattern produced by the following corner: $\theta_0 = 3\pi/2$ (e). The corner aperture is shown at left.

deemed unwarranted. As one might expect, increasing L improved the angular accuracy of the diffraction pattern representation; increasing P improved the radial accuracy.

To supplement the three-dimensional plots and to see easily the positions of the spikes, we plotted the logarithm of the normalized intensity vs theta for a fixed value of the radial coordinate in Fig. 3.8 (χ = 10). A relative maximum (at least one to two orders of magnitude above the background) corresponds to a bright or intensity spike in the diffraction pattern. The angular locations of these spikes agree with the description presented in connection with the three-dimensional plots; their profiles are illustrated clearly.

The intensity values plotted in Fig. 3.8 are normalized so that the intensity at the center of the diffraction pattern is unity. Such normalization is useful when we are examining the individual patterns. On the other hand, it is also beneficial to compare the relative strengths of the central diffraction peaks for different opening angles θ_0. In Table 3.1 we show how this peak changes as the opening angle varies. A clear aperture ($\theta_0 = 2\pi$) has been assigned a central diffraction peak intensity of 1.0000. The variation as a function of θ_0 is due to the differing areas of the apertures. If we assume that the amplitude of the illuminating electric field is the same for each aperture (hence, the same total optical power is incident on each aperture), then an aperture with two times the area transmits four times the optical power. Thus, the central diffraction peak of the π edge is four times greater than the peak of the $\pi/2$ corner and the peak of the $3\pi/2$ corner is nine times greater than the peak of the $\pi/2$
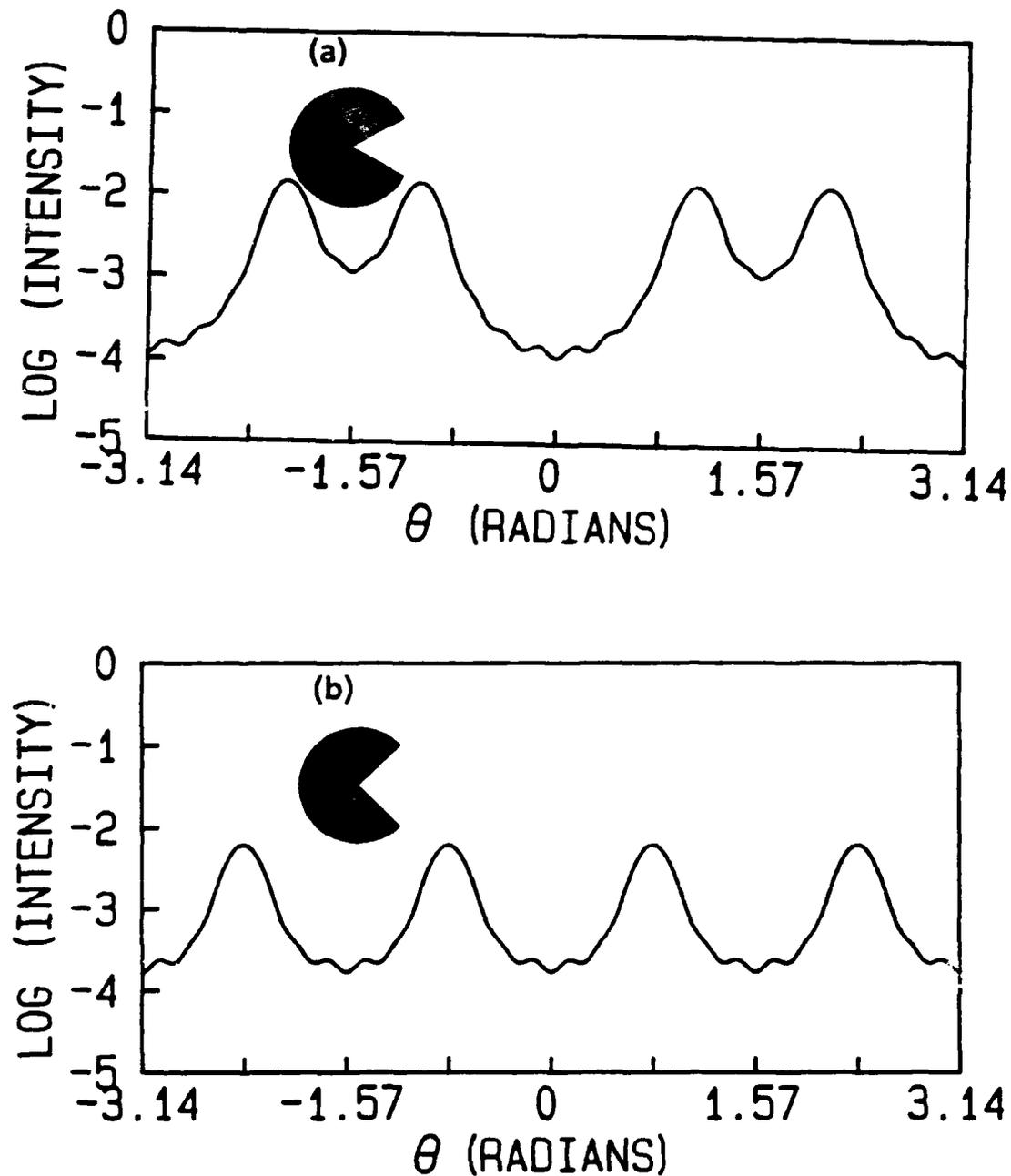
Fig 3.8. Variation of log(intensity) with theta for the corner diffraction pattern. Shown are GBE calculations for the logarithm of the normalized optical intensity of the diffraction pattern produced by the following corners: $\theta_0 = \pi/3$ (a), $\pi/2$ (b). The calculations were performed at a fixed radial coordinate, i.e., $\chi = 10$.
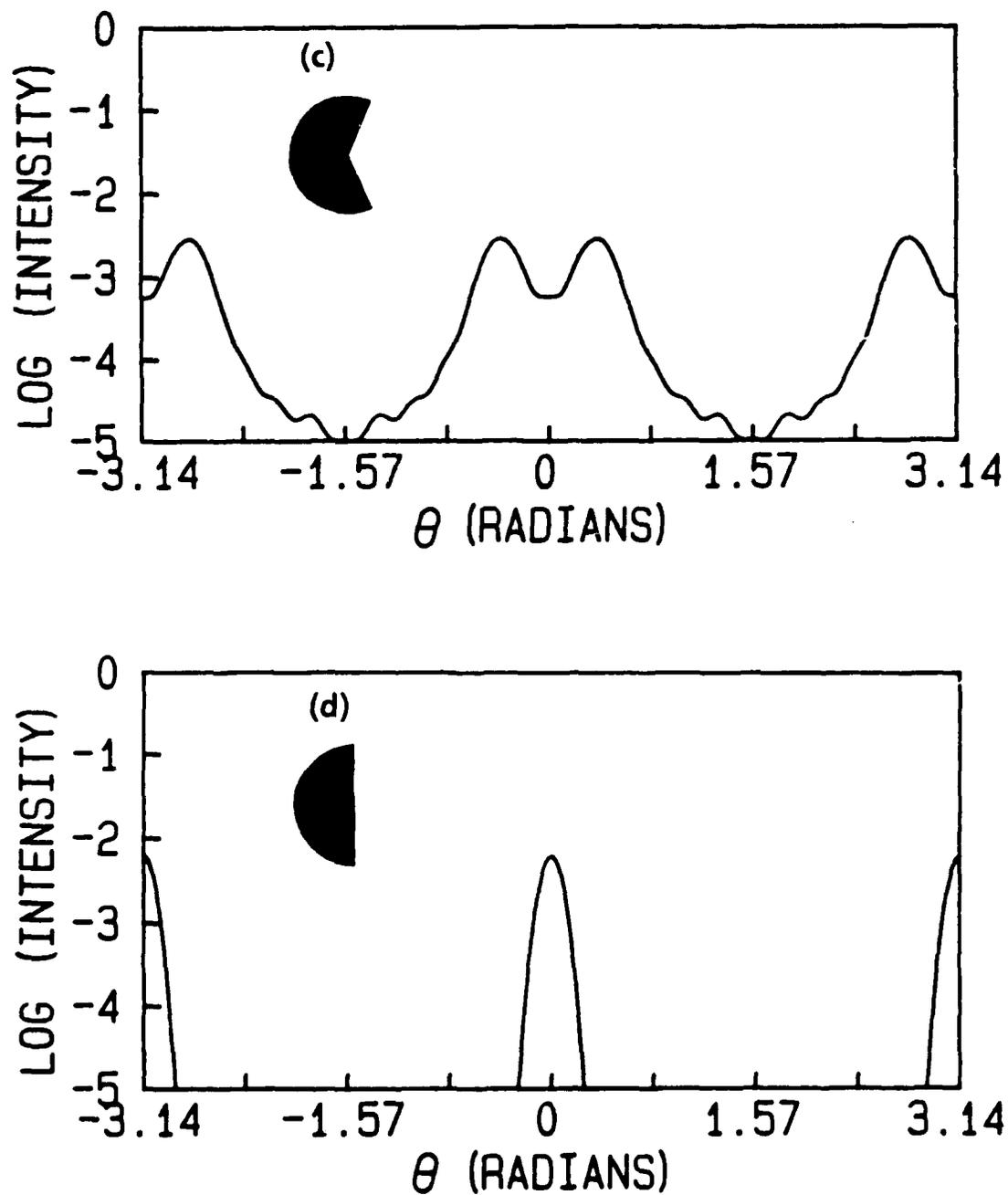
Fig. 3.8. Variation of log(intensity) with theta for the corner diffraction pattern. Shown are GBE calculations for the logarithm of the normalized optical intensity of the diffraction pattern produced by the following corners: $\theta_0 = 3\pi/4$ (c), π (d). The calculations were performed at a fixed radial coordinate, i.e., $\chi = 10$.
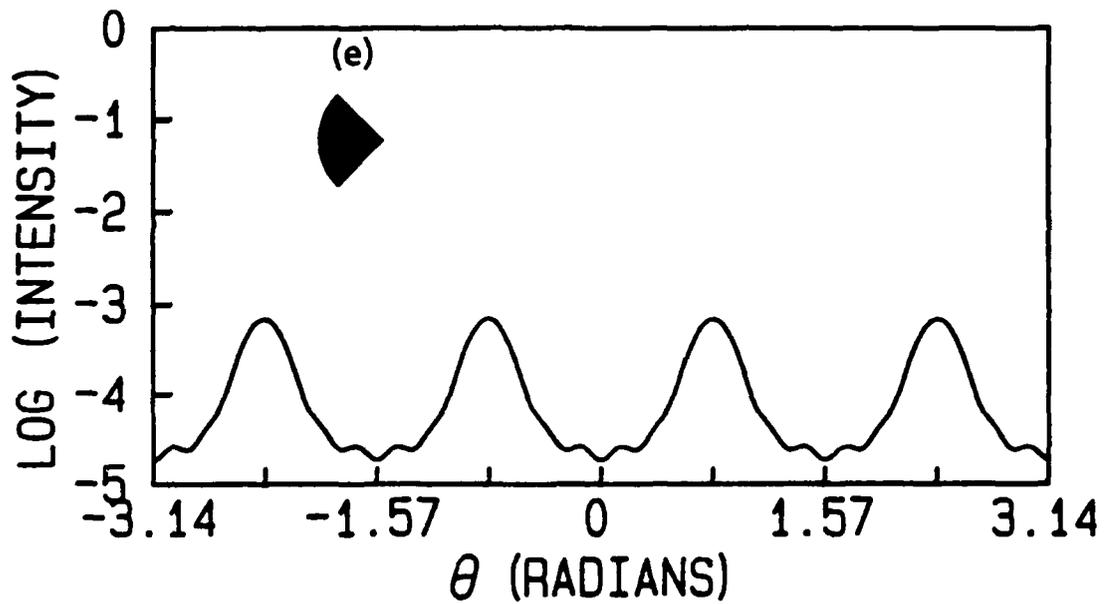
Fig. 3.8. Variation of log(intensity) with theta for the corner diffraction pattern. Shown is GBE calculation for the logarithm of the normalized optical intensity of the diffraction pattern produced by the following corner: $\theta_0 = 3\pi/2$ (e). The calculation was performed at a fixed radial coordinate, i.e., $\chi = 10$.

Table 3.1. Relative Corner Diffraction Peak
Intensities vs θ_0.

θ_0	Relative Diffraction Peak Intensity
$\pi/3$	0.0278
$\pi/2$	0.0625
$3\pi/4$	0.1406
π	0.2500
$3\pi/2$	0.5625

corner. For example, in Figs. 3.8(b) and (e), which are for the complementary apertures $\pi/2$ and $3\pi/2$, the only difference in the curves is a normalization factor of nine. We examine in greater detail later the similarities and differences of these two diffraction problems.

To check qualitatively the validity of the preceding results, photographs were taken of each of the corners. We present the photographs in Fig. 3.9. The photographs were recorded in the focal plane of a converging Gaussian beam: $w' \approx 5\text{mm}, -R' = z_0 \approx 1000\text{mm}$. The diffracting corners were masks of chrome on glass. The photographs were overexposed somewhat to enhance the appearance of the spikes. The orientation of the spikes agrees with the GBE predictions. Figure 3.9(a) shows approximately the distance corresponding to $\chi = 10$.

In Fig. 3.7(e) we observe a curious feature that is not found in the other four patterns: a hole or null in the intensity pattern. There are two nulls occurring at $(\chi_x, \chi_y) \approx (0, \pm 2.12)$. The nulls are not numerical artifacts. If one solves precisely the Fresnel diffraction integral for a plane-screen corner with opening angle $\theta_0 = 3\pi/2$, then the location of the
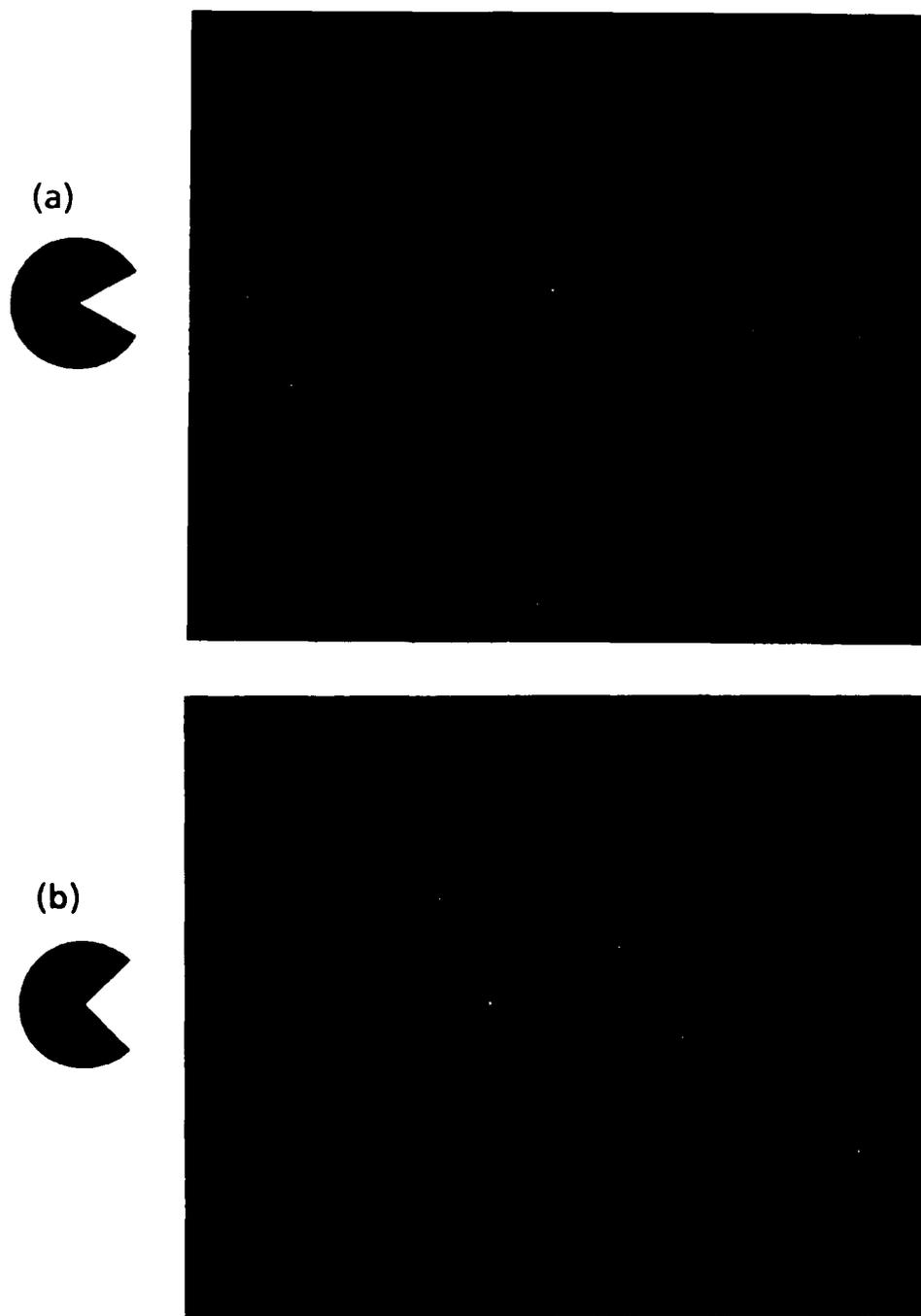
Fig. 3.9. Photographs of corner diffraction patterns for converging Gaussian illumination. The diffracting corners are represented at left; experimentally, they were chrome on glass masks. Shown are the patterns for $\theta_0 = \pi/3$ (a), $\pi/2$ (b). In (a) the distance $\chi = 10$ is shown.
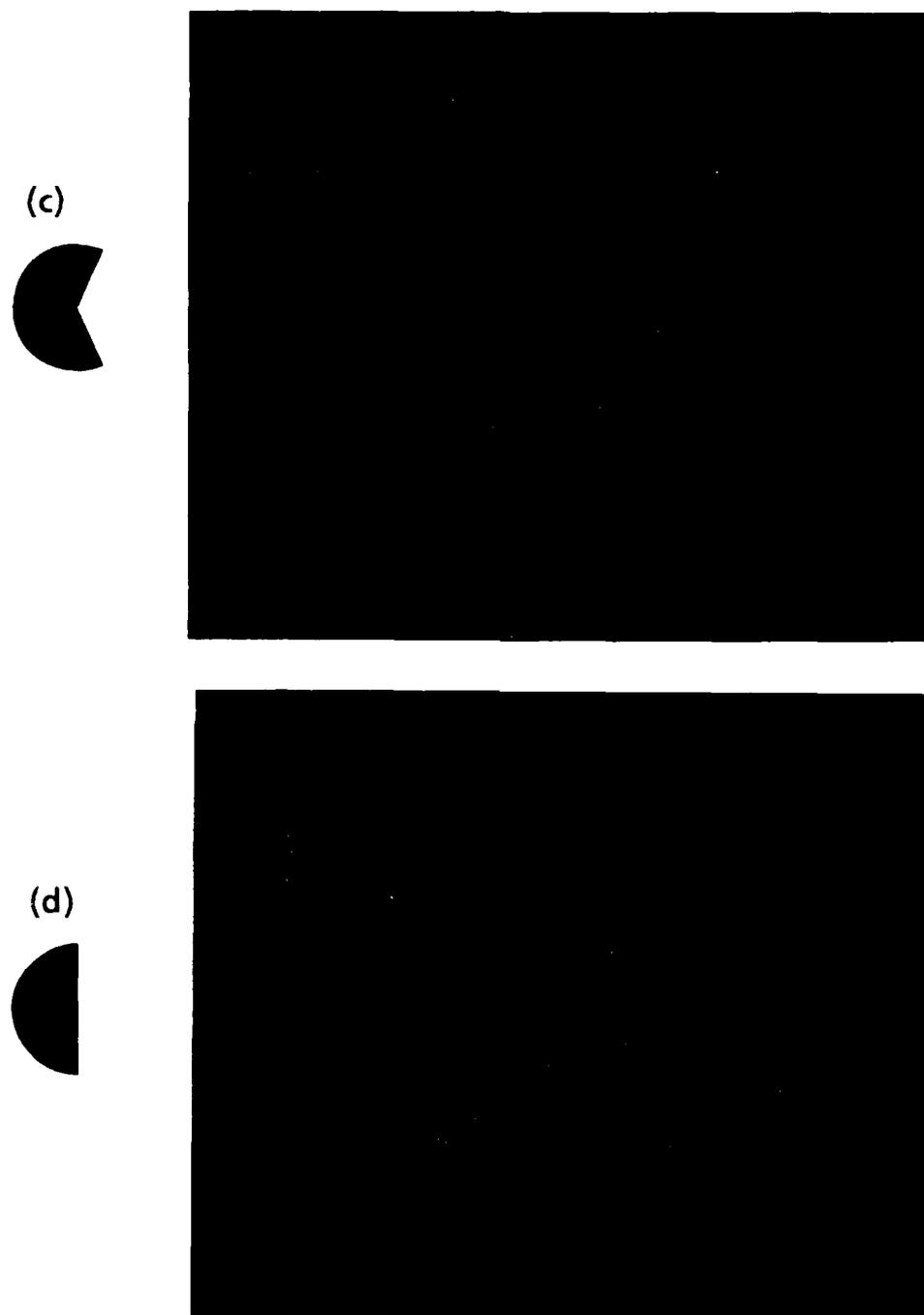
84

Fig. 3.9. Photographs of corner diffraction patterns for converging Gaussian illumination. The diffracting corners are represented at left; experimentally, they were chrome on glass masks. Shown are the patterns for $\theta_0 - 3\pi/4$ (c), π (d).

(e)
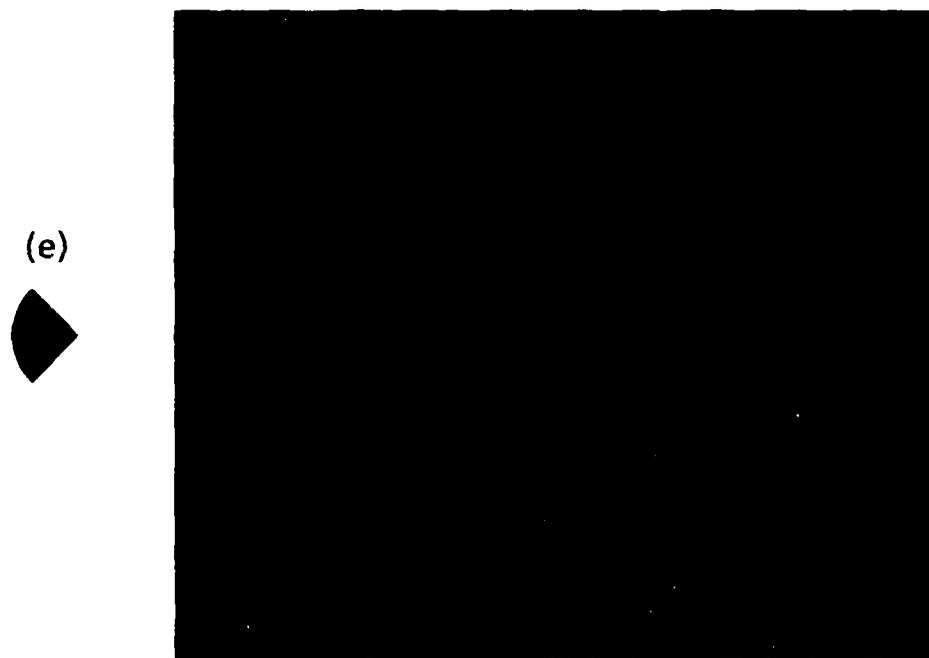
Fig. 3.9. Photograph of corner diffraction pattern for converging Gaussian illumination. The diffracting corner is represented at left; experimentally, it was a chrome on glass mask. Shown is the pattern for $\theta_0 = 3\pi/2$ (e).

nulls can be determined. The details of that calculation are not presented here, but the formula predicts that sharp nulls occur at $(\chi_x, \chi_y) = (0, \pm 2.0518)$. These points are located on the y axis in the observation plane at $y = \pm 1.4509 \, w$.

Further, these nulls are present for any plane-screen corner such that $\theta_0 > \pi$. Although we have not been able to write a closed-form expression for the diffraction integral arising from an opening angle not divisible by $\pi/2$, examination of GBE calculations for other large angles reveals similar intensity nulls. For example, in Fig. 3.10 we present a plot of the logarithm of the normalized intensity vs radial coordinate χ for a plane-screen corner with opening angle $\theta_0 = 5\pi/6$. The figure is log(intensity) vs radial coordinate χ for fixed θ ($\theta = \pi/2$), i.e., along the y axis. This corner result contains a strong null at $\chi \approx 2.07$.

To illustrate the dependence of this null position on opening angle θ_0, we present Fig. 3.11. Using the GBE solution, we determined the positions of the nulls for $\pi < \theta_0 < 2\pi$. Those positions are plotted vs θ_0 in the figure. We see that the null positions reach a minimum for $\theta_0 \approx 3\pi/2$. The nulls move out from the center as the opening angle approaches π (edge) or 2π (clear aperture).

We can gain some understanding of this phenomenon by comparing the diffraction patterns produced by a $\pi/2$ corner [Figs. 3.7(b) and 3.8(b)] and a $3\pi/2$ corner [Figs. 3.7(e) and 3.8(e)]. These two apertures are complements of each other, and we expect similar features to be present in both patterns. Indeed far enough from the center of the diffraction pattern (large χ), the patterns are the same. (In Figs. 3.7(b) and (e), the
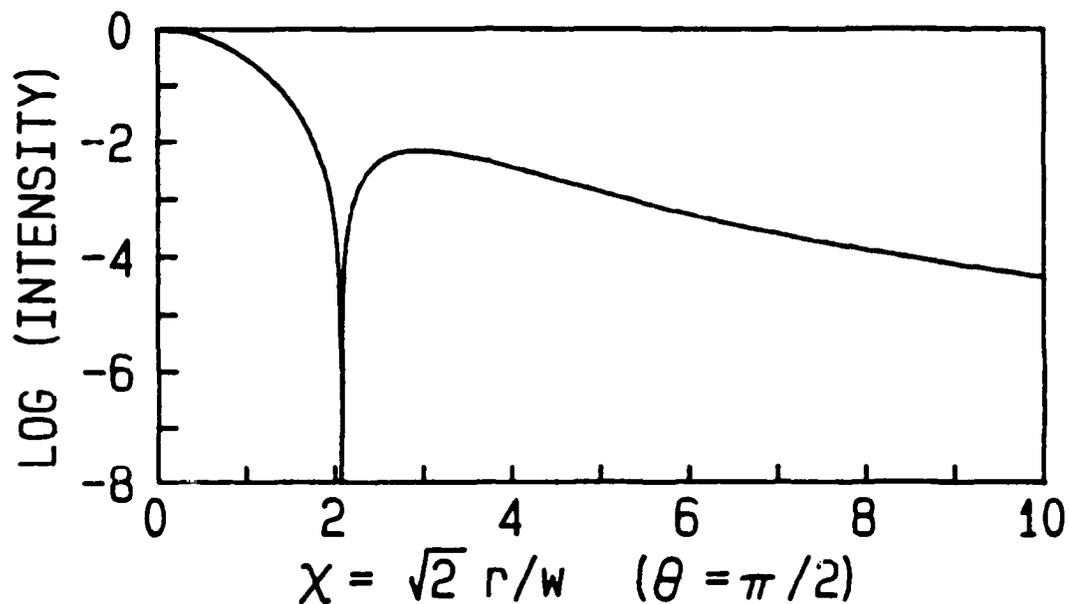
Fig. 3.10. Plot of log(intensity) along y axis for $\theta_0 = 5\pi/6$. The logarithm of the normalized intensity is plotted vs radial coordinate χ for fixed angular coordinate $\theta = \pi/2$. A sharp intensity null is seen at $\chi \approx 2.07$.
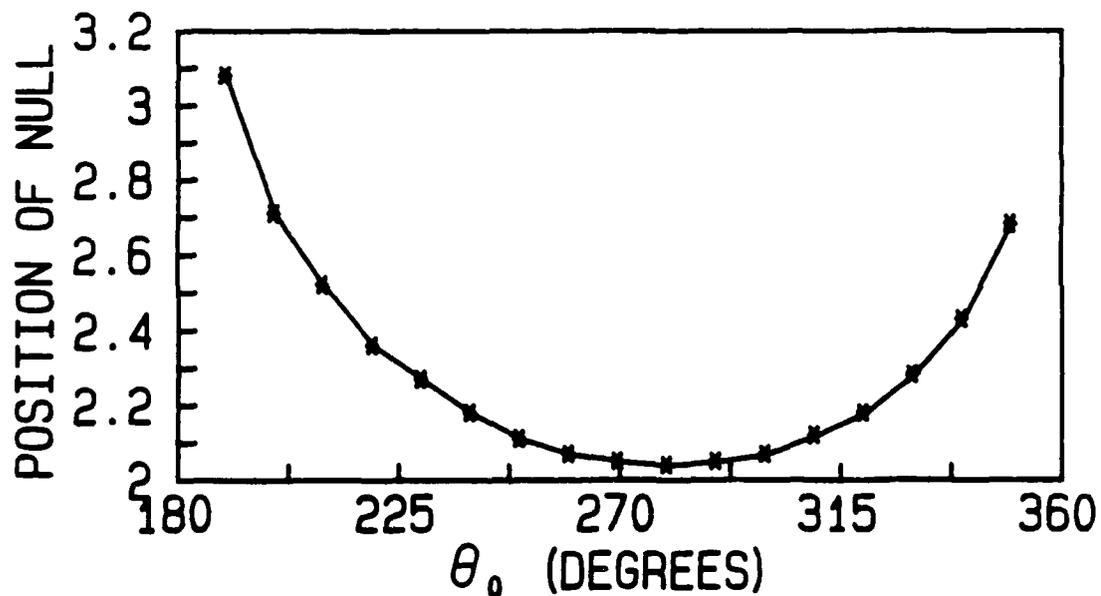
Fig. 3.11. Position of intensity nulls as a function of opening angle. The positions were determined from the GBE solution for a plane-screen corner of opening angle θ_0. Nulls occur for $\pi < \theta_0 < 2\pi$.

normalization factors for the patterns are different. We discussed these factors previously.) But, near the center of the diffraction pattern, the $3\pi/2$ pattern more closely resembles the pattern due to a clear aperture. By Babinet's principle the $3\pi/2$ pattern is the clear aperture pattern (unobstructed Gaussian beam) minus the $\pi/2$ pattern. One must combine these two patterns coherently. For small χ the clear aperture pattern dominates the nature of the $3\pi/2$ pattern; for large χ the $\pi/2$ contribution dominates. The intensity nulls appear in a transition region where the coherent subtraction of the $\pi/2$ corner pattern from the clear aperture pattern yields zero. Such a subtraction occurs for any opening angle such that $\theta_0 > \pi$.

Further validation of the GBE results is provided by comparison with a precise calculation of the modified Fresnel diffraction integral, Eq. (15), for an edge, $\theta_0 = \pi$. The published result (Livanos and George, 1975) is

$$e_x^{precise}(r,\theta,z_0) = \frac{\pi w'^2}{i\lambda z_0}\,exp\left[\frac{i\pi r^2}{\lambda z_0}\right]\left(\frac{1}{2}\right)exp\left[-\frac{y^2}{w^2}\right]\left\{1 - erf\left[\frac{ix}{w}\right]\right\} \quad . \tag{66}$$

(This is Eq. (27) of their paper rewritten to reflect an $exp(-i\omega t)$ time dependence and the notation of this chapter. Equation (66) is not written in the normalized radial coordinate.)

In Fig. 3.12 we compare a calculation from Eq. (66) with a calculation from Eq. (65), $\theta_0 = \pi$. The graph is a plot of $log(|e_x(\chi,\theta,z_0)|^2/|e_x(0,\theta,z_0)|^2)$ vs θ ($\chi = 10$). In other words, the plot is a scan over theta of the normalized intensity of the diffraction pattern at a fixed radial coordinate. The precise calculation (——) displays the single spike in intensity along $\theta = 0$ (and π). The GBE (- - - -) shows good agreement; especially near the spike, the two
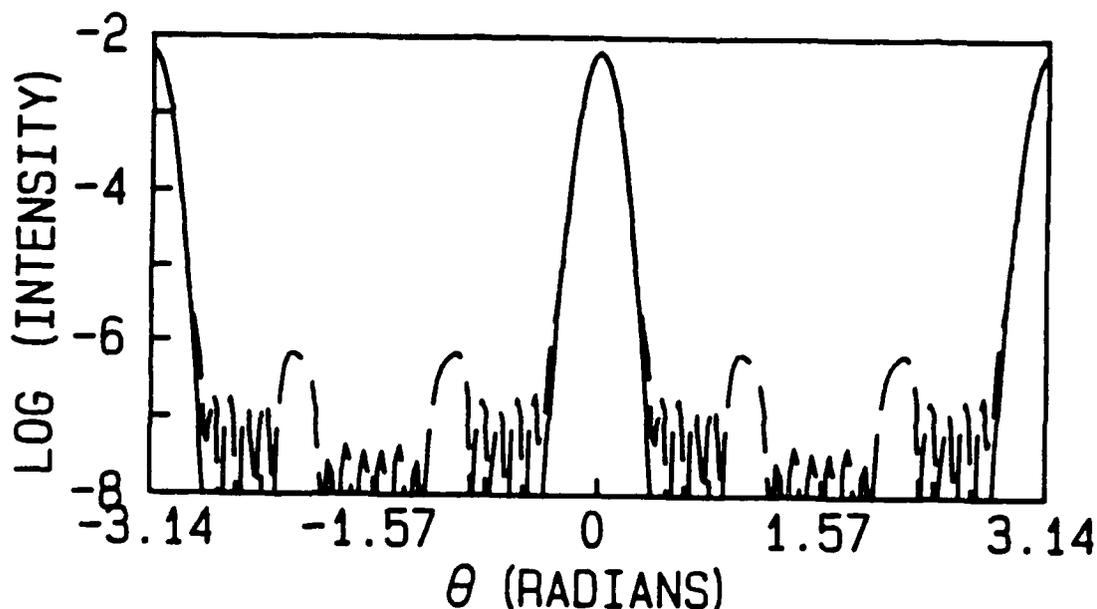
Fig. 3.12. Comparison of GBE representation of diffraction pattern of a corner of opening angle $\theta_0 = \pi$ (an edge) to precise published result. The logarithm of the normalized intensity is plotted vs theta for normalized radial coordinate $\chi = 10$. The solid line (———) represents the precise published result; the dashed curve (- - - -), is the GBE approximation.

curves are indistinguishable. In the dark region between the spikes, the precise formula is nearly zero ($< 10^{-10}$), but the GBE is not. We note, however, that the GBE calculation shown was computed for a finite number of terms ($P = 100$, $L = 25$). Increasing the angular resultion, i.e., increasing L, will improve the agreement between the GBE calculation and the precise formula, Eq. (66). For example, when $L = 35$, the ringing between the diffraction spikes is depressed to $< 10^{-9}$.

3.1.2.3   Gaussian Illuminated Double Corner

A small adjustment to the plane-screen corner GBE solution yields the solution to a related diffraction problem. Let the boundary value

function be

$$F(r',\theta') = \left[rect\left(\frac{\theta'}{\theta_0} \right) + rect\left(\frac{\theta'-\pi}{\theta_0} \right) \right] exp(-r'^2/w'^2)\, exp(i\pi r'^2/\lambda R') \quad . \qquad (67)$$

This aperture is a double corner: an opening of angle θ_0 centered at $\theta' = 0$ and an opening of angle θ_0 centered at $\theta = \pi$. The effect of the second term in the square brackets in Eq. (67) is to eliminate terms such that l is odd in Eq. (65); the contribution of even terms is doubled. The GBE solution to the boundary value problem described in Eq. (67) is

$$e_x^{DC}(\chi,\theta,z_0) = \frac{\pi w'^2}{i\lambda z_0} exp\left| \frac{i\lambda z_0 \chi^2}{2\pi w'^2} \right| 2 \sum_{p=0}^{P} \sum_{\substack{l=-L \\ even}}^{L} \left\{ \frac{\theta_0}{2\pi} \frac{sin[l\theta_0/2]}{l\theta_0/2} \frac{1}{(|l|+p)!} \left(\frac{|l|}{2} \right) \Gamma\left(p + \frac{|l|}{2} \right) \right\}$$

$$\times \chi^{|l|} L_p^{|l|}(\chi^2)\, exp(il\theta)\, exp(-\chi^2/2)\, (-i)^{2p+|l|} \qquad (68)$$

A three-dimensional representation of a GBE calculation for this double corner is presented in Fig. 3.13 ($\theta_0 = \pi/2$). The plot is similar to Fig. 3.7(b), which is the result for a single corner ($\theta_0 = \pi/2$); there are two intensity spikes. There is, however, an interesting difference. Down the center of each spike runs an intensity null. This corresponds to a dark line along the spikes. This phenomenon is clearly evident in Fig. 3.14 where we have plotted the logarithm of the normalized intensity vs theta at a fixed radial coordinate ($\chi = 10$). This result should be compared to the single corner result, Fig. 3.8(b). The intensity null down the center of each spike is obvious. The second symmetrically placed corner serves to change an intensity maximum into an intensity minimum.

For the Gaussian illuminated aperture, this phenomenon has not been mentioned in the literature. Hence, we felt it worthwhile to confirm this diffraction pattern feature experimentally. In Fig. 3.15 we show a

0

-10

-20 dB Scale

-30

-40

0.5
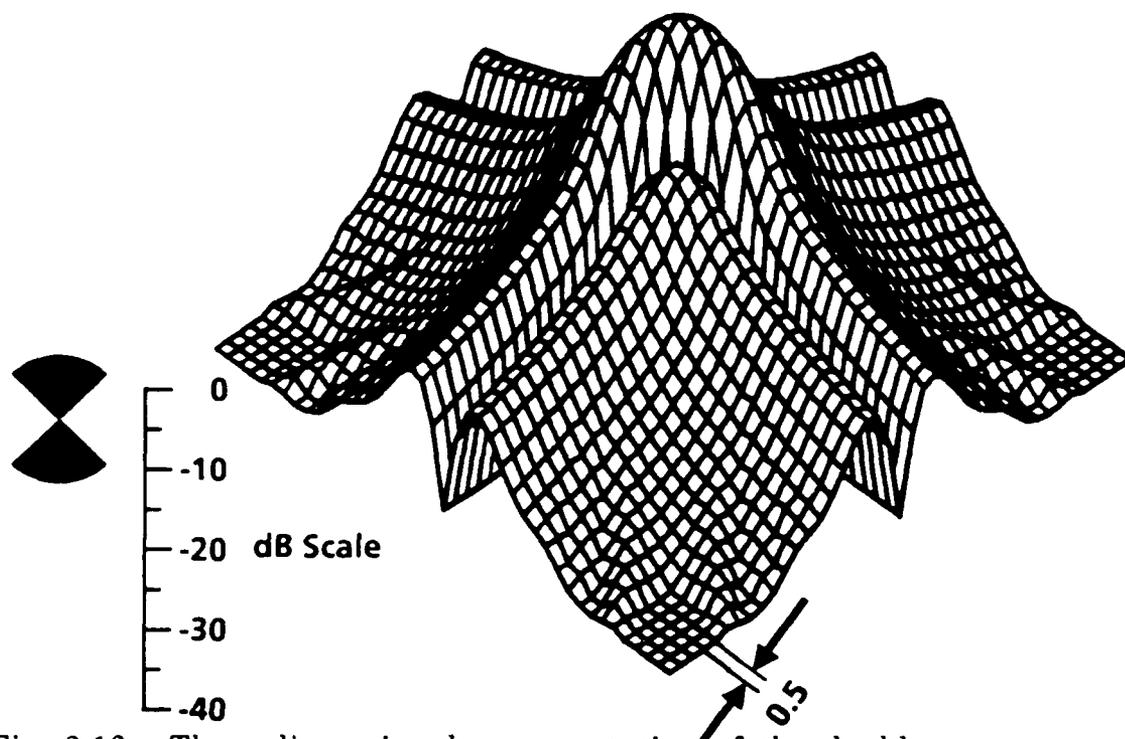
Fig. 3.13. Three-dimensional representation of the double corner diffraction pattern. Shown is GBE calculation for the logarithm of the normalized optical intensity of the diffraction pattern produced by a double corner ($\theta_0 = \pi/2$). The double corner aperture is shown at left.
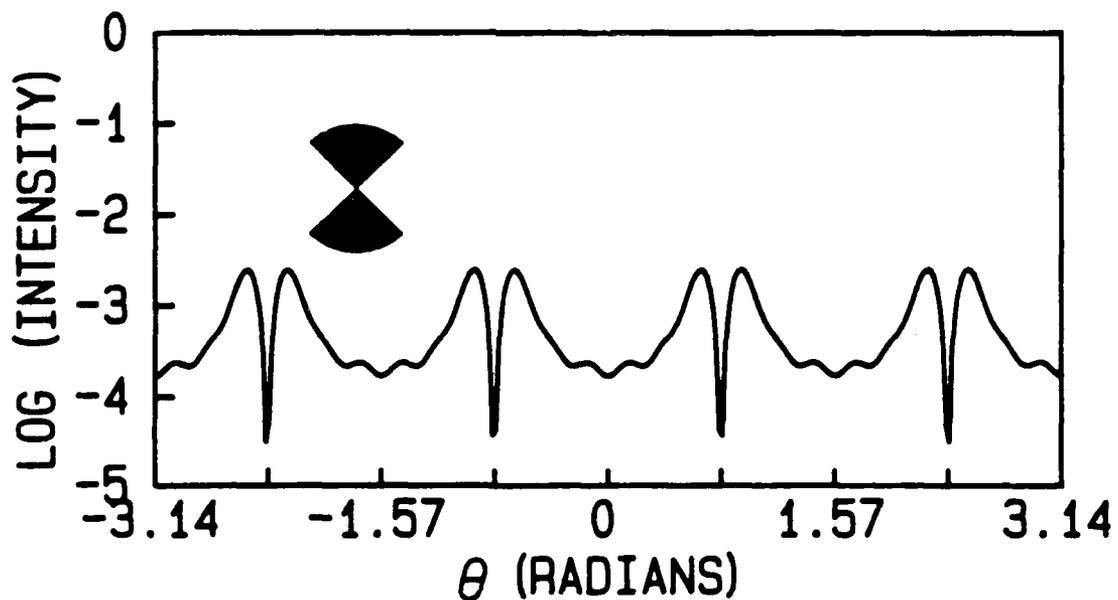
Fig. 3.14. Variation of log(intensity) with theta for the double corner diffraction pattern. Shown is GBE calculation for the logarithm of the normalized optical intensity of the diffraction pattern produced by a double corner ($\theta_0 = \pi/2$). The calculation was performed at a fixed radial coordinate, i.e., $\chi = 10$.

photograph of the pattern produced by a π/2 double corner. The pattern is characterized by two intensity spikes, which are perpendicular to the corresponding edge in the aperture. The intensity nulls down the center of the spikes are evident; thus, the GBE prediction is confirmed.


Fig. 3.15. Photograph of double corner diffraction pattern for converging Gaussian illumination. The diffracting corner is represented at left; experimentally, it was a chrome on glass mask. Shown is the pattern for a double corner with $\theta_0 = \pi/2$. The presence of an intensity null down the center of each spike is confirmed by the photograph.

3.2 Fourier Transform Analysis

Although the GBE has great flexibility in treating the corner diffraction problem, direct computation of the optical transform provides some additional insights and an interesting comparison to the GBE calculations. In addition, we use the calculations in the following section (§3.2.1) in Chapter 4 to describe diffraction from polygonal and nested polygonal apertures.

As a brief statement of how the Fourier transform arises in optical diffraction pattern analysis, we begin with Eq. (15), the modified Fresnel diffraction integral. If we assume the boundary value function $F(x',y')$ to have a quadratic phase factor corresponding to a converging wave, i.e.,

$$F(x',y') = f(x',y') \, exp\left[-\frac{i\pi}{\lambda z_0}\left(x'^2 + y'^2 \right) \right] \quad , \tag{69}$$

then the field in the focal plane $z = z_0$ is [using Eqs. (15) and (69)]

$$e_x(x,y,z_0) = \frac{1}{i\lambda z_0} \, exp\left[-\frac{i\pi}{\lambda z_0}\left(x^2 + y^2 \right) \right]$$

$$\times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x',y') \, exp\left[-\frac{i\pi}{\lambda z_0}(xx' + yy') \right] dx'dy' \quad . \tag{70}$$

Except for the complex phase factor in front of the integral, the field in the focal plane, $e_x(x,y,z_0)$, is a Fourier transform of $f(x',y')$ evaluated at spatial frequencies $u = x/\lambda z_0$ and $v = y/\lambda z_0$. The Fourier transform definition,

$$\bar{f}(u,v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x',y') \, exp[-i2\pi(ux' + vy')] dx'dy' \quad , \tag{71}$$

substituted into Eq. (70) gives

$$e_x(x,y,z_0) = \frac{1}{i\lambda z_0} \, exp\left[-\frac{i\pi}{\lambda z_0}\left(x^2 + y^2 \right) \right] \bar{f}\left(\frac{x}{\lambda z_0}, \frac{y}{\lambda z_0} \right) \quad . \tag{72}$$

The use of the bar here denotes Fourier transform. Hence, determination

of the Fourier transform of the modified boundary value function $f(x',y')$ solves the diffraction problem.

3.2.1   Corner Bounded by a Slit Aperture

Consider a plane-screen corner of bounded extent, namely, a plane-screen corner bounded by the function $rect(x'/2C)$; this aperture is also identified as a triangular aperture, a triangular section of a circular aperture, or simply a section (see Fig. 3.16). For this aperture

$$f^{section}(x',y') = \begin{cases} 1 & -ax' \leq y' \leq ax' \ and \ 0 \leq x' \leq C \\ 0 & otherwise \end{cases} \tag{73}$$

This aperture is related to the original description of the plane-screen corner, i.e., $rect(\theta'/\theta_0)$ in Fig. 3.1, by setting $a = tan(\theta_0/2)$.
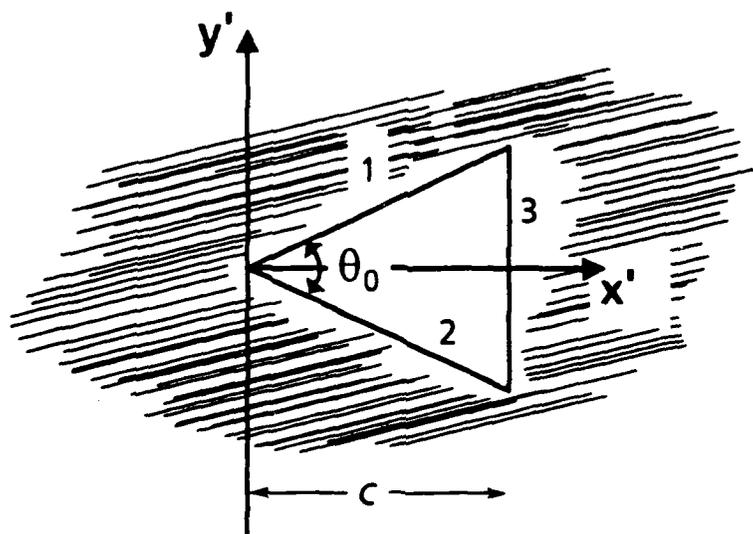

Fig. 3.16. Corner bounded by a slit aperture (also known as a triangular or section aperture). The opening angle is θ_0, and the corner is bounded by edge 3 at $x = C$, i.e., the corner is an isosceles triangle with height or altitude $= C$. The three edges are labeled 1, 2, and 3.

The calculation of the Fourier transform follows.

$$\bar{f}^{section}(u,v) = \int_0^C dx' \int_{-ax'}^{ax'} dy'\ exp[-i2\pi(ux' + vy')]\quad,$$

$$\bar{f}^{section}(u,v) = \int_0^C dx'\left(-\frac{1}{i2\pi v}\right)[\ exp(-i2\pi w_+ x') - exp(-i2\pi w_- x')]\quad, \tag{74}$$

where

$$w_\pm = u \pm av \quad. \tag{75}$$

The final result is

$$\bar{f}^{section}(u,v) = \left(-\frac{C}{i2\pi v}\right)[\ exp(-i2\pi C w_+)\ sinc(Cw_+) - exp(-i2\pi C w_-)\ sinc(Cw_-)] \tag{76}$$

Equation (76) agrees with published results (see, for example, Sillitto and Sillitto, 1975).

Before understanding the characteristics and features of this transform, we investigate the behavior of $\bar{f}^{section}$ near $u = 0$ and $v = 0$. Consider first the limiting case $u \neq 0, v \to 0$:

$$\lim_{\substack{u \neq 0 \\ v \to 0}} \bar{f}^{section}(u,v) = \left(-\frac{C}{i2\pi}\right) exp(-i\pi Cu)$$

$$\times \lim_{\substack{u \neq 0 \\ v \to 0}} \frac{exp(-i2\pi Cav)\ sinc[C(u+av)] - exp(+i2\pi Cav)\ sinc[C(u-av)]}{v} \quad.$$

Since the numerator and denominator both give zero for $v = 0$, L'Hospital's rule for determining the limit is an appropriate technique to employ. We take the derivative with respect to v in both the numerator and denominator and tries to evaluate at $v = 0$. Proceeding in such a manner,

$$\lim_{\substack{u \neq 0 \\ v \to 0}} \bar{f}^{section}(u,v) = (aC^2)\ exp(-i\pi Cu)\left\{ sinc(Cu) + \frac{i}{\pi Cu}[cos(\pi Cu) - sinc(Cu)]\right\} \quad. \tag{77}$$

Hence, the leading factor in Eq. (76), which contains $1/v$, does not pose any difficulty; the result near the u axis is finite.

Now consider $u \to 0$, $v \neq 0$:

$$\lim_{\substack{u \to 0 \\ v \neq 0}} \bar{f}^{section}(u,v) = (aC^2)\, sinc^2(Cav) \quad . \tag{78}$$

Again this limit is finite.

Finally, using either Eq. (77) or Eq. (78), we evaluate the center value, i.e., $(u,v) = (0,0)$:

$$\lim_{\substack{u \to 0 \\ v \to 0}} \bar{f}^{section}(u,v) = (aC^2) \quad . \tag{79}$$

Now we can write the transform normalized to unity at the origin as

$$\bar{f}_0^{section}(u,v) = \left(\frac{i}{2nCav} \right) [\, exp(-i2nCw_+)\, sinc(Cw_+) - exp(-i2nCw_-)\, sinc(Cw_-)\,] \tag{80}$$

To understand the features of this Fourier transform or the related diffraction pattern, examine Fig. 3.17. The logarithm of the square modulus of Eq. (80) is computed and displayed as an 8-bit gray-scale image. White is high intensity; black is low intensity. We scaled the computed data so that the image represents seven orders of magnitude; the scaling is described by the following formula:

$$pixel\ value = \begin{cases} \left| \dfrac{log(|\bar{f}_0^{section}(u,v)|^2)}{7} + 1 \right| \times 255 & 10^{-7} < |\bar{f}_0^{section}(u,v)|^2 < 10^0 \\[2ex] 0 & |\bar{f}_0^{section}(u,v)|^2 < 10^{-7} \end{cases}$$

Table 3.2 shows some sample mappings of intensity to pixel value. The parameter values used to compute Fig. 3.17 were $C = 2.24$ mm, $\theta_0 = 72°$ (a = 0.7265), $\lambda = 0.6328$ μm, and $z_0 = 444$ mm. Each pixel corresponds to a 50 μm × 50 μm area in the optical transform plane. Since the image is 512 × 512 pixels, the observation plane coordinates (x,y) fall in the range (–12.8

mm, 12.8 mm). The frequency variables $(u,v) = (x/\lambda z_0, y/\lambda z_0)$ fall in the range $(-45.56 \text{ mm}^{-1}, 45.56 \text{ mm}^{-1})$.

Clearly the three intensity spikes dominate the pattern in Fig. 3.17. The spike along $\theta = 0°$ is perpendicular to edge 3 in Fig. 3.16. The other two spikes are perpendicular to edge 1 (along $\theta = 36° + 90°$) and edge 2 (along $\theta = -36° + 90°$). Edges cause spikes in the optical transform. The half-width of a spike, i.e., the distance from the relative maximum along the spike to the first minimum measured perpendicular to the spike, is equal to the reciprocal of the length of the edge that caused it. These
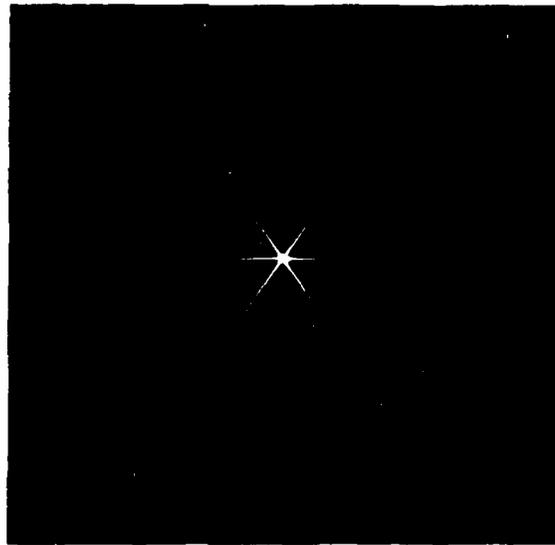
Fig. 3.17. Gray-scale representation of section diffraction pattern. The logarithm of the square modulus of Eq. (80), $|f_0^{section}(u,v)|^2$ is displayed as an 8-bit gray-scale image. The mapping of the logarithm of the intensity to pixel value is linear such that an intensity of unity yields pixel value 255 and an intensity below 10^{-7} yields pixel value 0. Details of the intensity mapping are presented in the text. The results shown here are for a section with height $C = 2.24$ mm, $\theta_0 = 72°$. The transform configuration is $\lambda = 0.6328$ μm and $z_0 = 444$ mm. The image is 512×512 pixels with each pixel representing an area of 50 μm × 50 μm.

Table 3.2. Sample Intensity to Pixel Value
Mappings.

Normalized Intensity Value, e.g., $\lvert f_0^{section}(u,v) \rvert^2$	Pixel Value
1.0000	255
10^{-1}	219
5.6×10^{-2}	209
10^{-3}	146
10^{-5}	73
3.24×10^{-6}	55
10^{-7}	0
$< 10^{-7}$	0

widths are given by the following relations:

$$edge\,1: \quad width\,of\,spike = \left\lvert C(1+a^2)^{1/2} \right\rvert^{-1} \quad ,$$

$$edge\,2: \quad width\,of\,spike = \left\lvert C(1+a^2)^{1/2} \right\rvert^{-1} \quad ,$$

$$edge\,3: \quad width\,of\,spike = \left\lvert 2aC \right\rvert^{-1} \quad .$$

An examination of the behavior of the transform along the center of the spikes reveals that the field drops off like $1/r$ where r is the distance from the center of the transform; the drop-off is monotonic. The drop-off of the field along a radial line not coincident with a spike is $1/r^2$.

To understand better the section diffraction pattern, we present Fig. 3.18. In these plots we show the dependence of the section diffraction pattern intensity in a plane perpendicular to the z axis at a fixed radial coordinate in that plane, $r = (u^2 + v^2)^{1/2} = 1.5$ mm. The logarithm of the
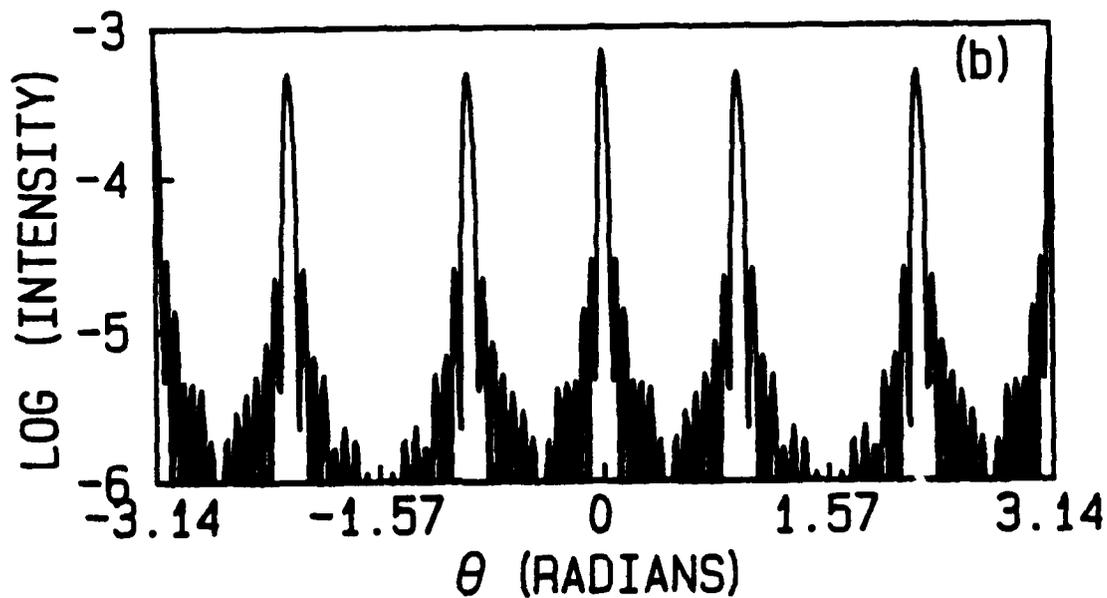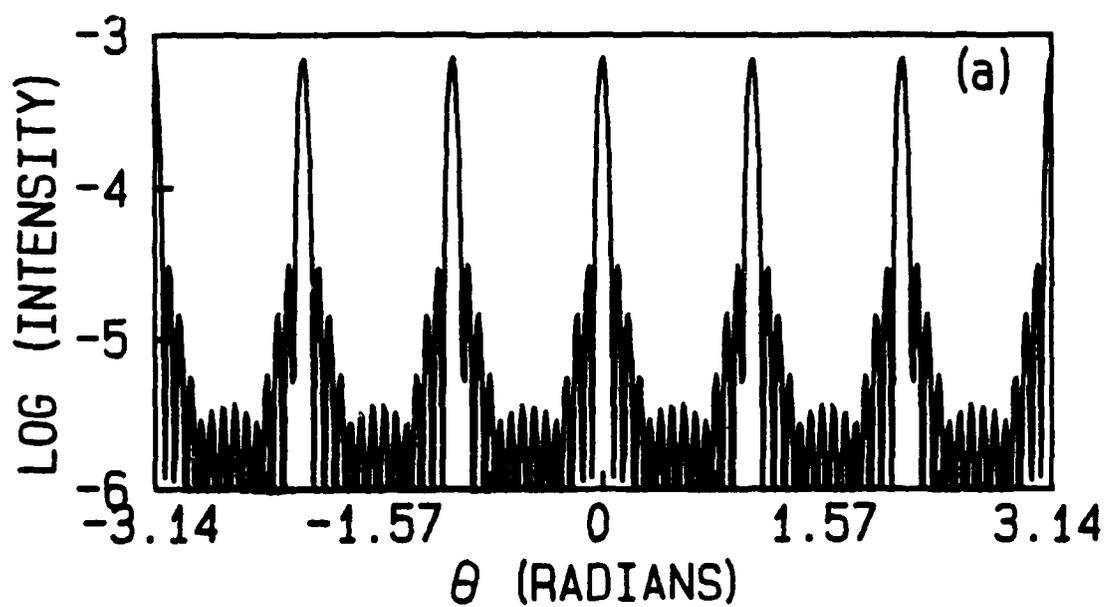
Fig. 3.18. Angular scans of log(intensity) for various section apertures. The radial coordinate is fixed at $r = 1.5$ mm, and the log(intensity) is plotted vs θ for (a) $\theta_0 = 60°$ and (b) $\theta_0 = 72°$. The plots show how the position of the spikes and their widths depend on the opening angle.
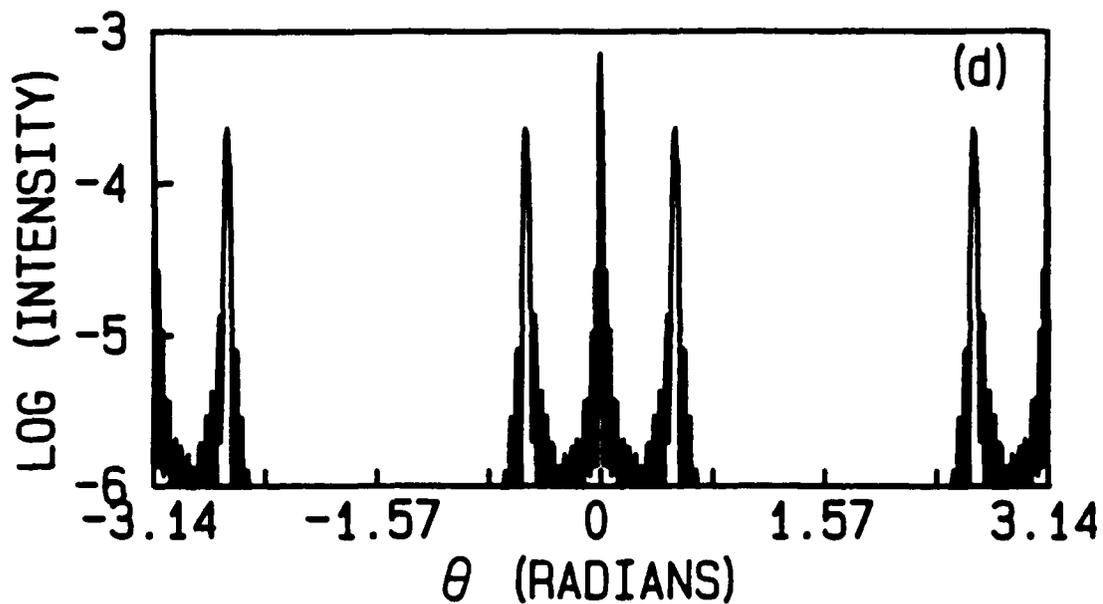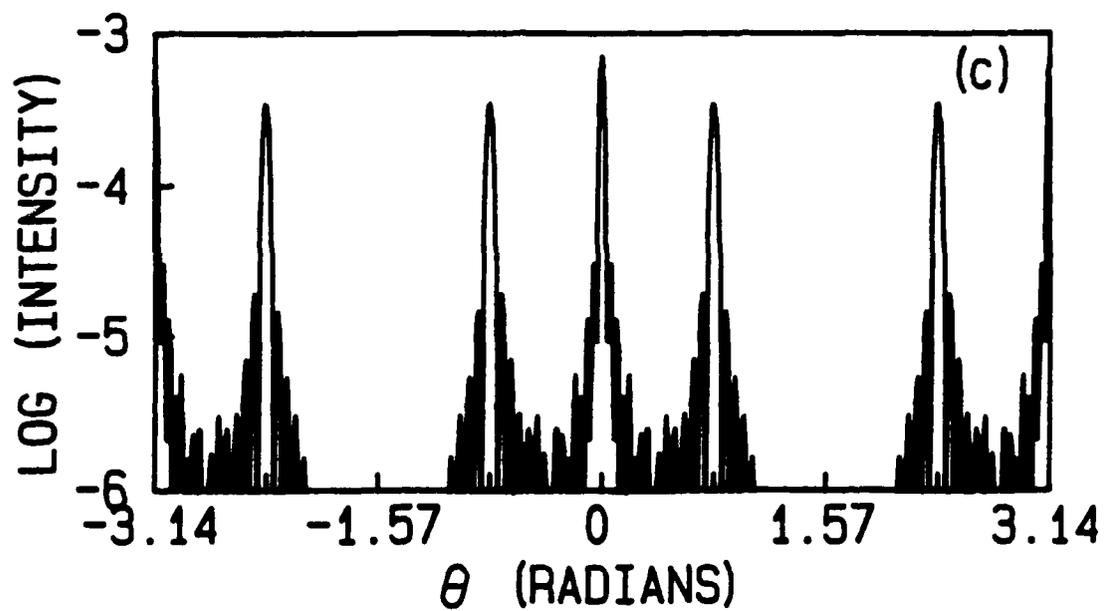
Fig. 3.18. Angular scans of log(intensity) for various section apertures. The radial coordinate is fixed at $r = 1.5$ mm, and the log(intensity) is plotted vs θ for (c) $\theta_0 = 90°$ and (d) $\theta_0 = 120°$. The plots show how the position of the spikes and their widths depend on the opening angle.

normalized intensity vs the angle θ plots show the spikes, which are visible in Fig. 3.17, as local maxima. The spikes occur at the angular locations specified previously. Four opening angles, $\theta_0 = 60°, 72°, 90°$, and $120°$, are displayed to illustrate the phenomenon. Notice that for $\theta_0 = 60°$, all of the spikes are of equal width and all of the subsidiary maxima are of the same height. This aperture is an equilateral triangle. Thus, the spikes are equispaced in the transform plane, and the symmetry of the aperture is preserved. For the case of $\theta_0 = 120°$, edge 3 is 1.73 times longer that edges 1 and 2; as a result, this spike is narrower by the same factor.

3.2.2 Corner Bounded by a Circular Aperture

Finally, we consider a sector of angle θ_0 as shown in Fig. 3.19. This aperture is described by $rect(\theta'/\theta_0)circ(r'/a)$ or

$$f^{sector}(x',y') = \begin{cases} 1 & -ax' \le y' \le ax' \ and \ 0 \le (x'^2+y'^2)^{1/2} \le a \\ 0 & otherwise \end{cases} \quad (81)$$

where $a = tan(\theta_0/2)$.

The calculation of the Fourier transform follows.

$$\bar{f}^{sector}(u,v) = \int_0^C dx' \int_{-ax'}^{ax'} dy' \ exp[-i2\pi(ux' + vy')]$$

$$+ \int_C^a dx' \int_{-(a^2-x'^2)^{1/2}}^{(a^2-x'^2)^{1/2}} dy' \ exp[-i2\pi(ux' + vy')] \quad , \quad (82)$$

where

$$C = a \cos(\theta_0/2) = \frac{a}{(1+a^2)^{1/2}} \quad . \quad (83)$$

We recognize the first term in Eq. (82) as the Fourier transform of a section; this calculation was performed in §3.2.1. Hence, we can write
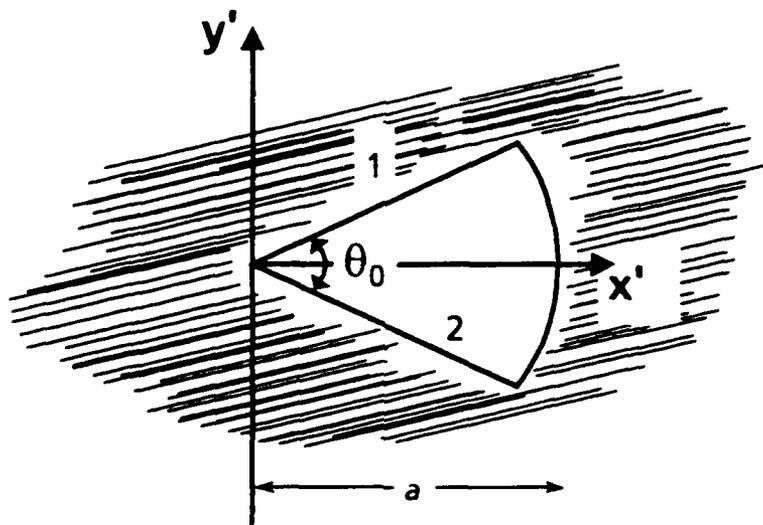
Fig. 3.19. Corner bounded by a circular aperture or sector of a circular aperture. The opening angle is θ_0, and the corner is bounded by a circle of radius a and edges 1 and 2.

$$\bar{f}^{sector}(u,v) = \bar{f}^{section}(u,v) + \bar{g}(u,v) \quad , \tag{84}$$

where

$$\bar{g}(u,v) = \int_C^a dx' \int_{-(a^2-x'^2)^{1/2}}^{(a^2-x'^2)^{1/2}} dy' \, exp[-i2\pi(ux' + vy')] \quad . \tag{85}$$

Equation (76) is the expression for $\bar{f}^{section}$ with C given by Eq. (83). The term $\bar{g}$ requires calculation.

First, we integrate $\bar{g}$ with respect to y'. Then we expand the integrand in Bessel functions that can be integrated term-by-term. The general result obtained, Eqs. (88) and (89), is well suited to numerical calculations.

Integration with respect to y' gives

$$\bar{g}(u,v) = \int_C^a dx' \, exp[-i2\pi(ux')] \left(\frac{1}{\pi v}\right) sin\left|2\pi v(a^2-x'^2)^{1/2}\right| \quad . \tag{86}$$

Substitution of $x' = a\cos(\theta')$ in Eq. (86) gives

$$\overline{g}(u,v) = \int_{\theta_0/2}^{0} \sin(\theta')\,d\theta'\left(-\frac{a}{nv}\right)exp[-i2nau\cos(\theta')]\sin[2nav\sin(\theta')] \quad . \tag{87}$$

Using the Bessel identities from Abramowitz and Stegun (1970, 9.1.43–45) and integrating gives the following result.

$$\overline{g}(u,v) = (na^2)\left|\frac{2}{\pi(2nav)}\right|$$

$$\times \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\left(1-\frac{\delta_{0m}}{2}\right)(-i)^m J_m(2nau)J_{2n+1}(2nav)\,g_1(m,2n+1) \quad , \tag{88}$$

where

$$g_1(\mu,v) = \frac{\sin[(1+\mu-v)\theta_0/2]}{(1+\mu-v)} - \frac{\sin[(1+\mu+v)\theta_0/2]}{(1+\mu+v)}$$

$$+ \frac{\sin[(1-\mu-v)\theta_0/2]}{(1-\mu-v)} - \frac{\sin[(1-\mu+v)\theta_0/2]}{(1-\mu+v)} \quad . \tag{89}$$

We can compute the Fourier transform of a sector of opening angle θ_0, Eq. (84), via Eqs. (75), (76), (83), (88), and (89).

Figure 3.20 contains angular scans of the sector diffraction pattern. In these plots we show the dependence of the sector diffraction pattern intensity in a plane perpendicular to the z axis at a fixed radial coordinate in that plane, $r = (u^2+v^2)^{1/2} = 1.5$ mm. The figure should be compared with Fig. 3.18 for the section diffraction patterns. The curves are normalized so that the intensity at $(u,v) = (0,0)$ is unity. There are two intensity spikes apparent. The spikes are produced by edges 1 and 2. Since edge 3 was replaced by a circular arc, there is no strong spike along the direction $\theta = 0°$; however, the intensity is non-zero along this direction. Four opening angles, $\theta_0 = 60°$, $72°$, $90°$, and $120°$, are displayed to illustrate
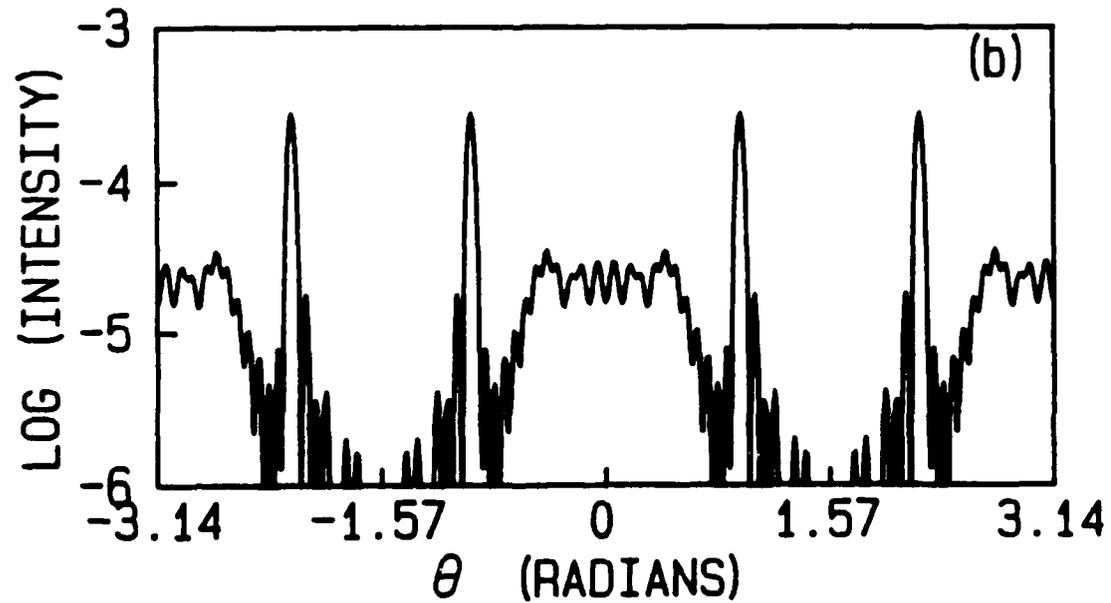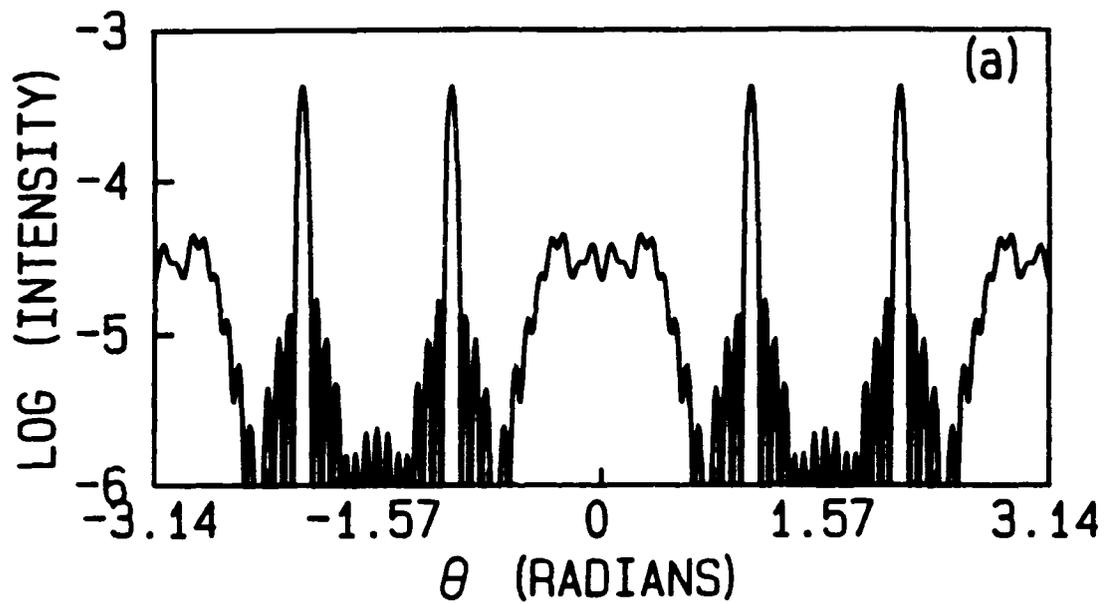
Fig. 3.20. Angular scans of log(intensity) for various sector apertures. The radius is fixed at $r = 1.5$ mm, and the log(intensity) is plotted vs θ for (a) θ_0 = 60° and (b) θ_0 = 72°. The following parameters were used: $a = 2.77$ mm, $\lambda = 0.6328$ μm, and $z_0 = 444$ mm.
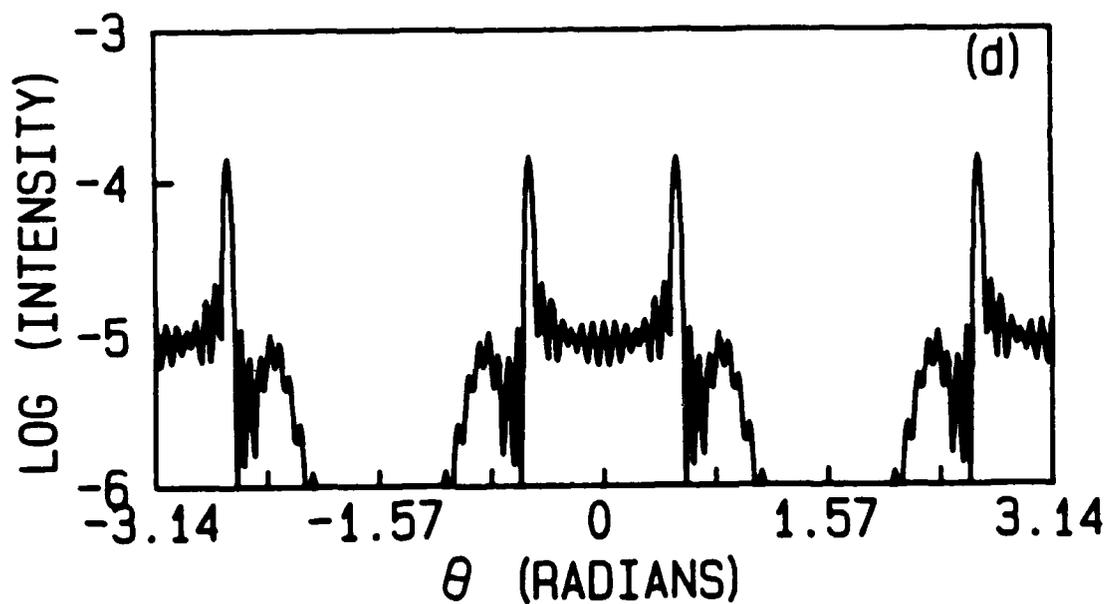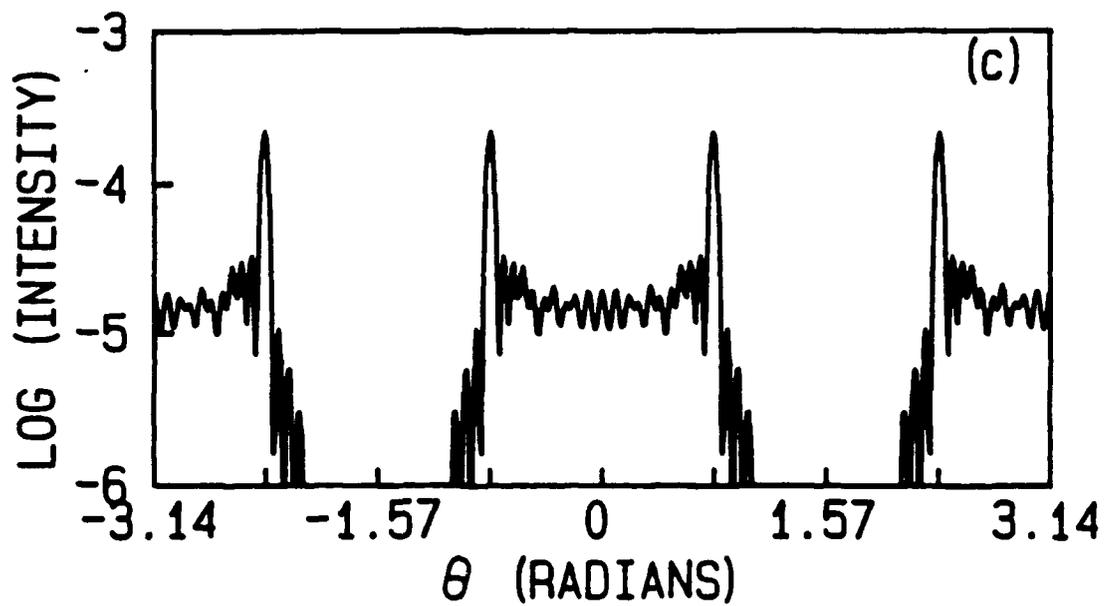
Fig. 3.20. Angular scans of log(intensity) for various sector apertures. The radius is fixed at $r = 1.5$ mm, and the log(intensity) is plotted vs θ for (c) θ_0 = 90° and (d) θ_0 = 120°. The following parameters were used: $a = 2.77$ mm, $\lambda = 0.6328$ µm, and $z_0 = 444$ mm.

the phenomenon. The parameters for the aperture and transform configuration are: $a = 2.77$ mm, $\lambda = 0.6328$ μm, and $z_0 = 444$ mm. Notice that for all opening angles, the spikes are of the same height and width because the edges producing them are equal.

Figures 3.21–3.24 are radial scans of transform patterns along $\theta = 0°$ (a) and along one of the spikes (b). The solid line (———) is the result for a circular aperture, i.e., the transform pattern is the Airy disk pattern. The first dashed curve (- - - -) is the result for the sector using Eq. (87). The second dashed curve (– - – -) is the result for a section using Eq. (76). All of the curves are normalized so that the intensity of each diffraction pattern is unity at $r = 0$. Notice that along $\theta = 0°$ (a), the section curve exhibits a stronger spike character than the sector curve. But the spike character is approximately the same along one of the spikes produced by edge 1 or 2. In addition, one observes that the sector curve is not monotonic as is the section curve.

With Fig. 3.25 we conclude the discussion of the sector transform pattern. The opening angle here is $\theta_0 = 180°$; the aperture is a semi-circle of radius a. The scan is along $\theta_0 = 0°$, which corresponds to the intensity spike in the pattern. The intensity along the spike oscillates in contrast to the intensity along the spike from a Gaussian illuminated edge [see Fig. 3.7(d)].

This semi-circular aperture calculation was performed by Livanos and George (1975). To see that the formulas presented here agree with their result, we must separate Eq. (87) into real and imaginary parts. Evaluating Eq. (88) for $\theta_0 = \pi$ and using Gradshteyn and Ryzhik (1980,
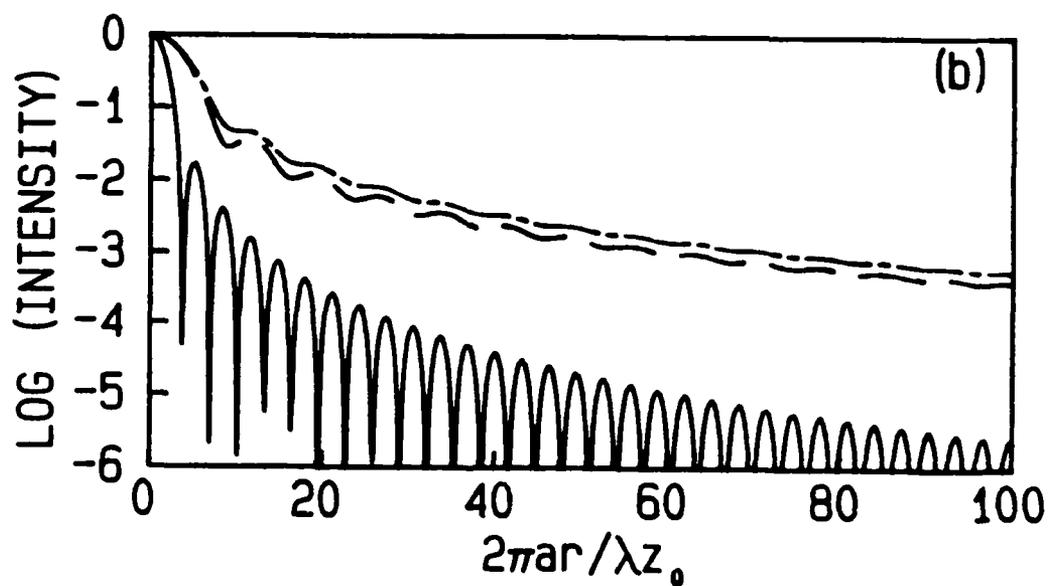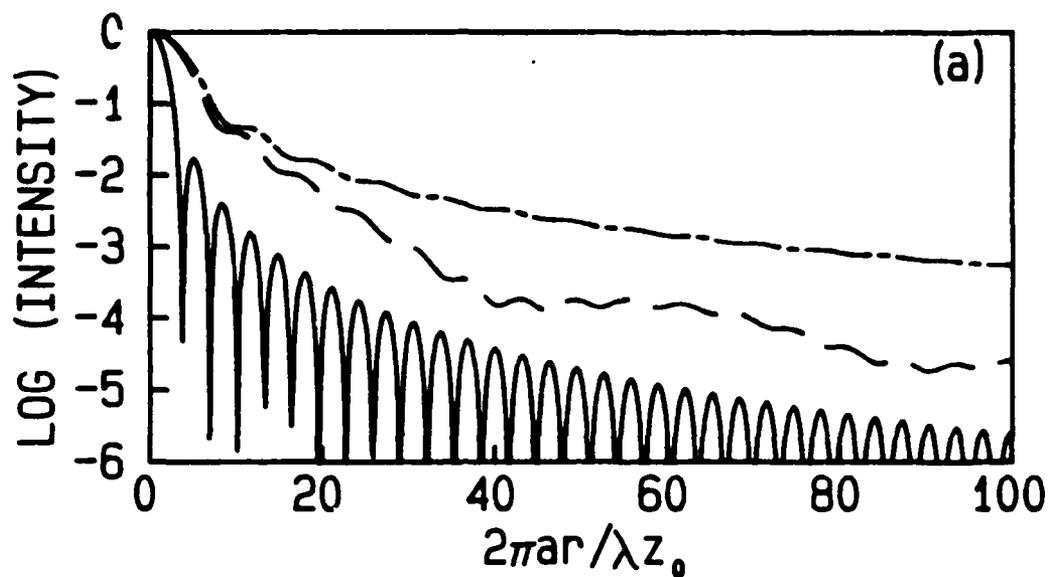
Fig. 3.21. Radial scans of log(intensity) for three apertures. The opening angle is $\theta_0 = 60°$. The scan is along (a) $\theta = 0°$ and (b) $\theta = 60°$. The three curves in each plot are for a circular aperture (———), a sector (- - - -), and a section (– - – -).

108
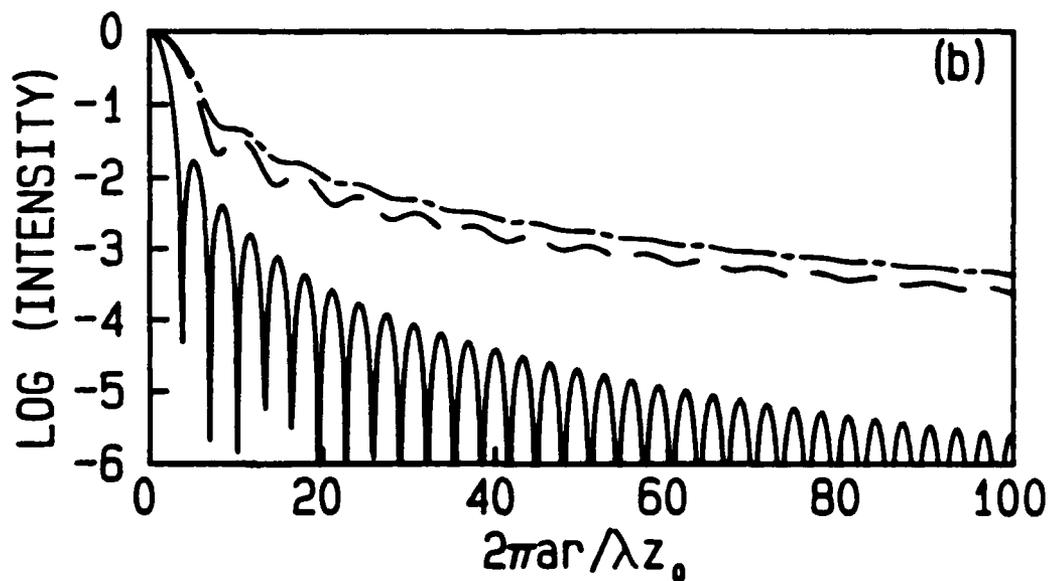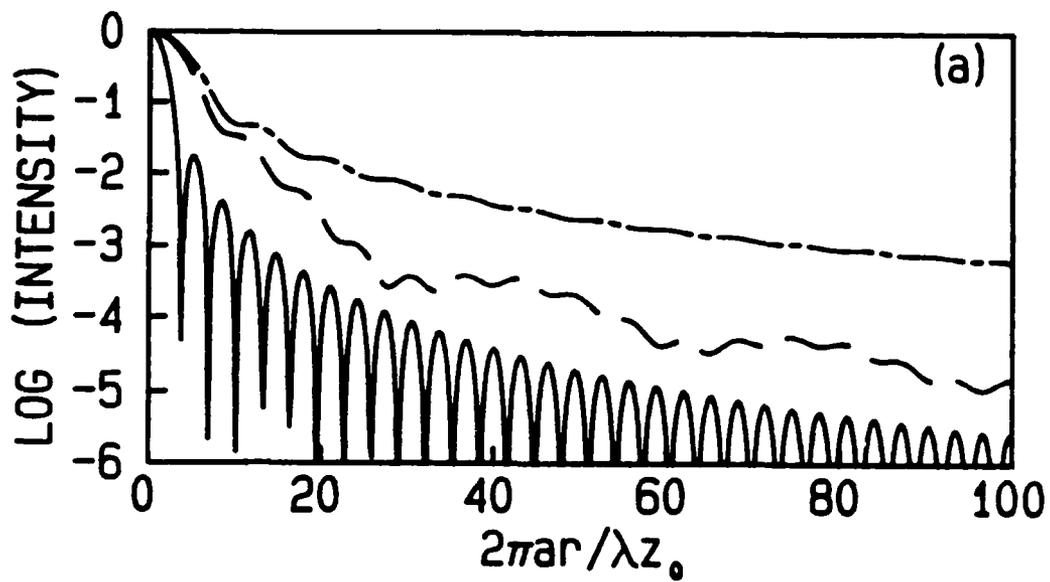
Fig. 3.22. Radial scans of log(intensity) for three apertures. The opening
angle is $\theta_0 = 72°$. The scan is along (a) $\theta = 0°$ and (b) $\theta = 54°$. The three
curves in each plot are for a circular aperture (———), a sector (- - - -), and a
section (– - – -).
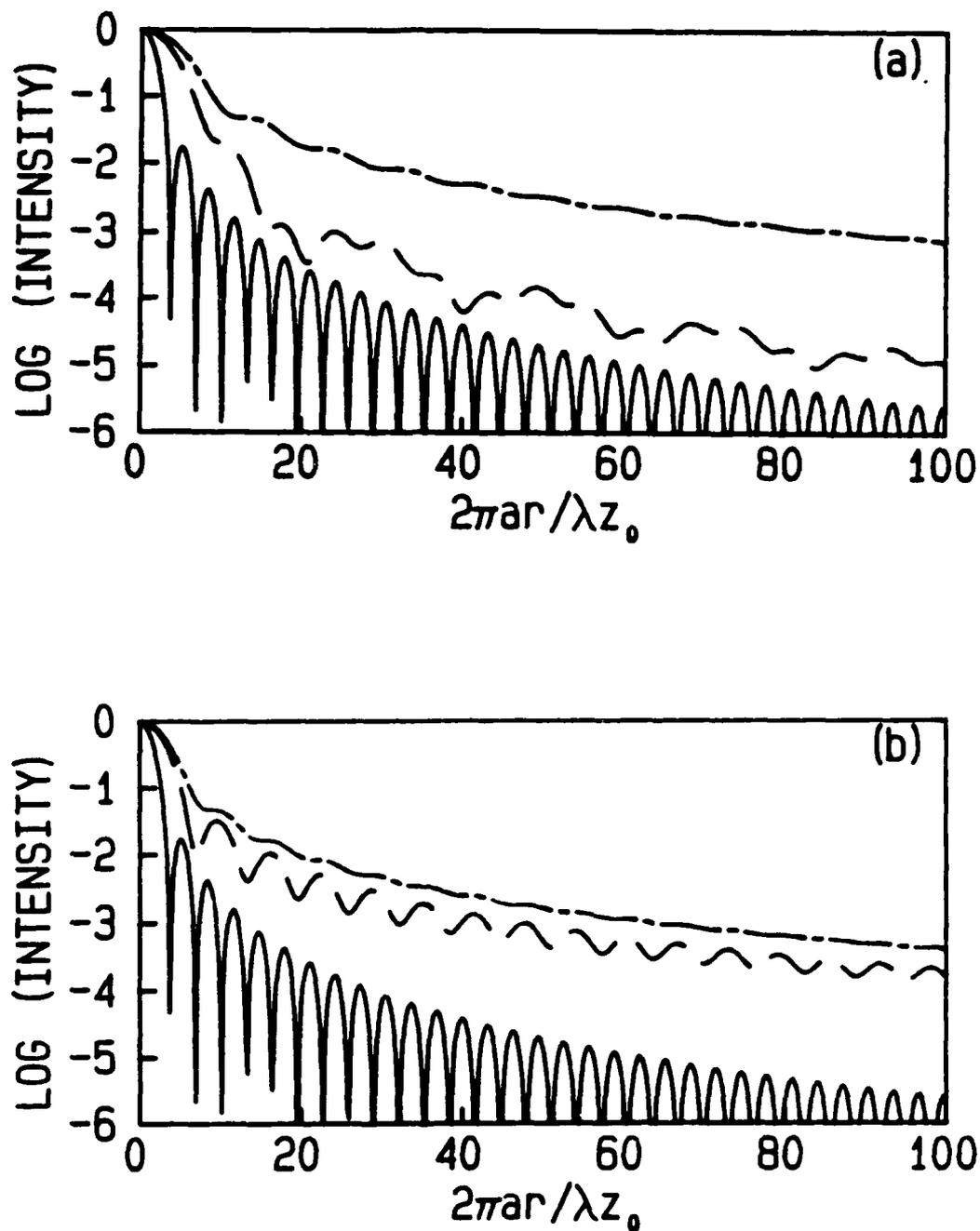
Fig. 3.23. Radial scans of log(intensity) for three apertures. The opening angle is $\theta_0 = 90°$. The scan is along (a) $\theta = 0°$ and (b) $\theta = 45°$. The three curves in each plot are for a circular aperture (——), a sector (- - - -), and a section (– - – -).
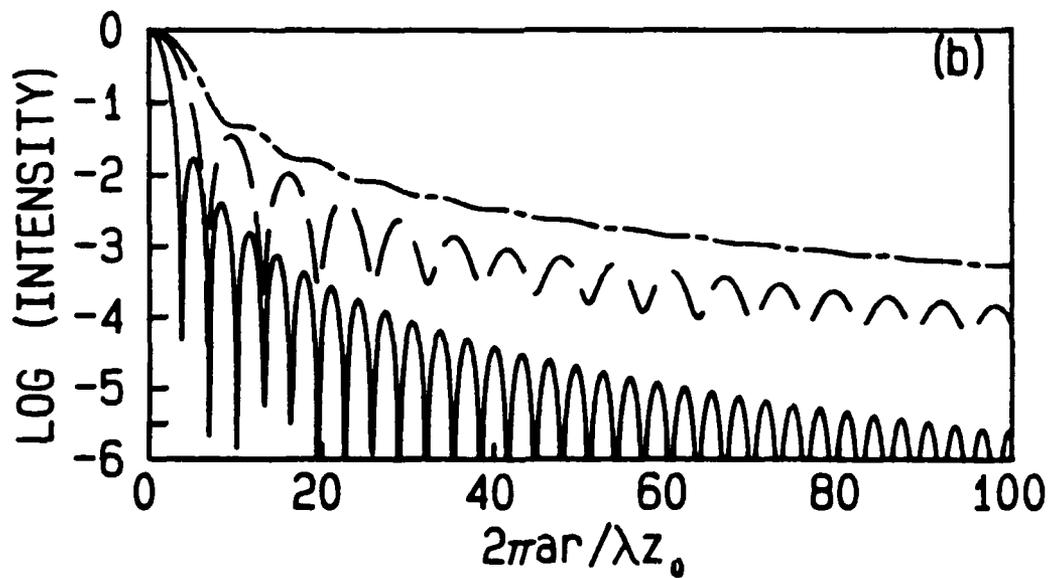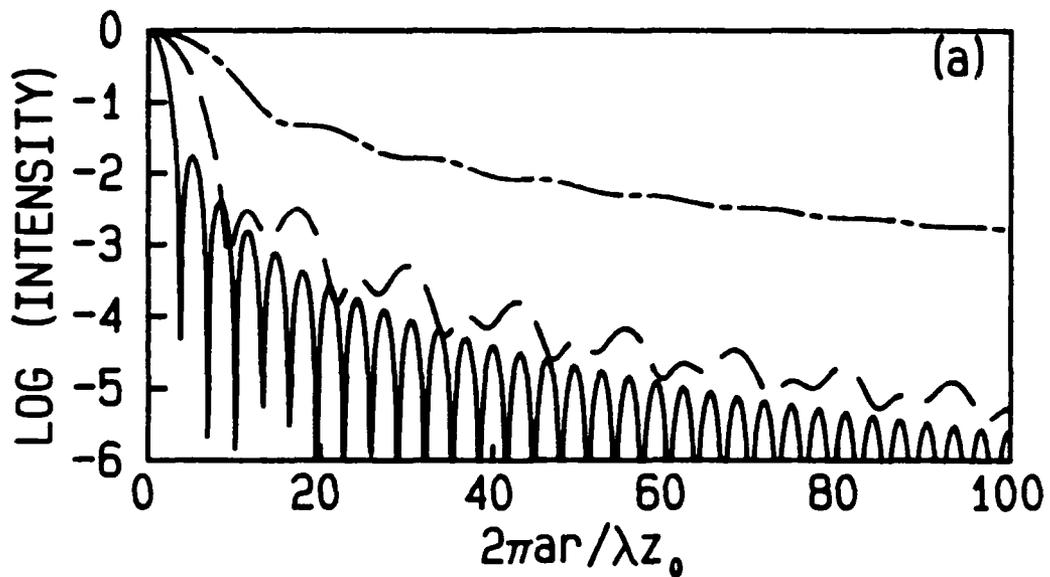
Fig. 3.24. Radial scans of log(intensity) for three apertures. The opening angle is $\theta_0 = 120°$. The scan is along (a) $\theta = 0°$ and (b) $\theta = 30°$. The three curves in each plot are for a circular aperture (——), a sector (- - - -), and a section (– · – ·).
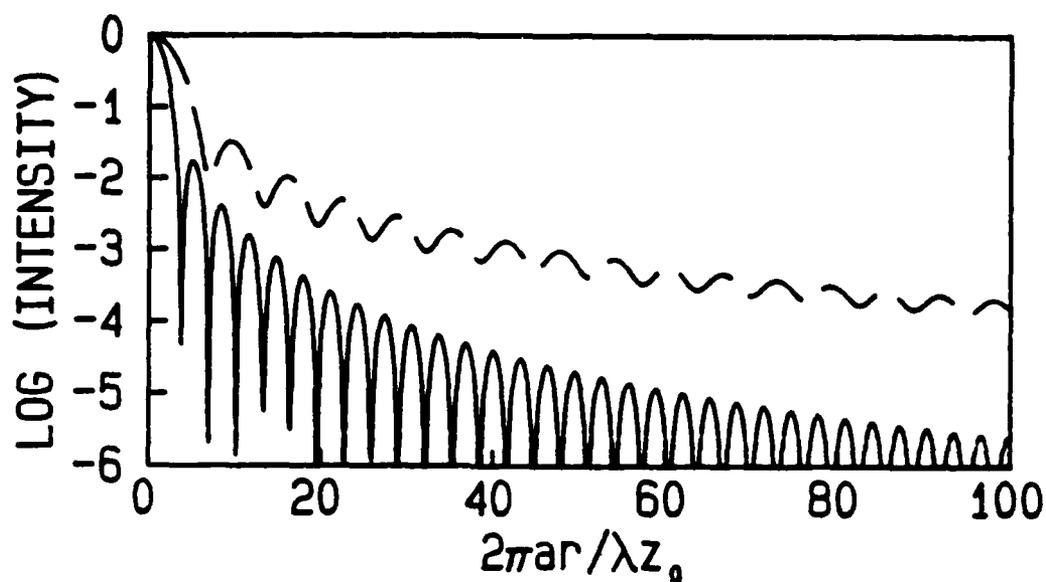
Fig. 3.25. Radial scans of log(intensity) for two apertures. The opening angle is $\theta_0 = 180°$. The scan is along $\theta = 0°$, i.e., along the intensity spike. The two curves are for a circular aperture (——) and a sector (- - - -). The sector aperture is a semi-circle.

8.53), one can show that the real part of Eq. (87) equals one-half times the

Airy disk function, i.e., Eq. (9) of their paper, and the imaginary part

equals Eq. (13) of their paper. In other words, the two formulas are

identical. Figure 3.25 is the same as Livanos and George's Fig. 3 except for

normalization.

3.3 Summary

In this chapter we studied diffraction by plane-screen corners. A

general plane-screen corner or opening of arbitrary angle θ_0 was

illustrated in Fig. 3.1. We performed all of the diffraction calculations

within the Fresnel approximation.

To treat the specific problem of a plane-screen corner illuminated by a Gaussian beam, we followed a familiar path in deriving the Fresnel diffraction integral, Eq. (13). A slight modification to this diffraction integral revealed an underlying boundary value problem [Eqs. (16) and (17)]. As an alternative to evaluating explicitly the diffraction integral, we described an eigenfunction approach: the Gaussian beam expansion [Eqs. (22), (24), and (26)]. In this approach one must solve for the expansion coefficients; for an arbitrary Gaussian beam illuminating a plane-screen corner of arbitrary angle, we presented the coefficients [Eq. (57)].

Computer calculations for the optical intensity pattern in the focal plane of a converging elementary Gaussian beam illuminating a plane-screen corner were performed for several opening angles. We showed three-dimensional plots of the normalized intensity (Fig. 3.7), plots of the logarithm of the normalized intensity vs angular coordinate (Fig. 3.8), and slightly overexposed photographs of the diffraction patterns (Fig. 3.9). The photographs confirmed the existence of features predicted by the computed results: intensity ridges or spikes that are produced by the edges in the aperture.

The accuracy of these results was established by a comparison to the published precise evaluation of the Fresnel diffraction integral for the case of an edge (see Fig. 3.12 and discussion). As one would expect, the accuracy of the expansion in approximating the precise solution depends on how many terms in the infinite summation are retained. On the other hand, the Gaussian beam expansion solution to this diffraction problem can be used for any opening angle $0 < \theta_0 < 2\pi$; we have not been able to evaluate

precisely the diffraction integral unless the opening angle is divisible by n/2.

The utility of the Gaussian beam expansion was demonstrated in diffraction pattern calculations for opening angles greater than n. The expansion predicts intriguing intensity nulls in the diffraction pattern. Since the Gaussian beam expansion solution is valid for arbitrary opening angle, we were able to determine the positions of these nulls as a function of opening angle (Fig. 3.11).

By means of a slight modification to the general solution, we obtained a solution to a different problem: diffraction of a Gaussian beam by a double corner. We described the aperture transmission function in Eq. (67), and we presented the solution in Eq. (68). The diffraction pattern for this aperture contains two perpendicular intensity spikes that have an intensity null down the center. The Gaussian beam expansion predicted this phenomenon (Figs. 3.13 and 3.14), and experimental observation confirmed it (Fig. 3.15).

To conclude our study of diffraction by plane-screen corners, we computed directly corner diffraction patterns produced in the focal plane of a uniform, converging beam. Straightforward evaluation of the Fourier transform of the aperture transmission function permitted us to compute these optical transform patterns.

In one case the corner was bounded by a vertical edge, thereby creating a triangular aperture; we usually referred to this aperture as a triangular section of a circular aperture or simply a section aperture (Fig. 3.16). Diffraction by this aperture will be a fundamental part of the

114

derivations presented in Chapter 4. We presented the diffraction solution
in Eq. (76).

The predicted optical transform pattern (Fig. 3.17), computed from
this formula, contains three intensity spikes; each edge in the section
aperture produces a spike. We represented the transform pattern as an 8-
bit gray-scale image in which the intensity data were scaled so that seven
orders of magnitude could be seen. Plots of the logarithm of the normalized
intensity vs angular coordinate at a fixed radius showed how the positions
and heights of the spikes change as a function of the opening angle (Fig.
3.18).

The second case considered was a corner bounded by a circular arc,
i.e., a sector of a circular aperture (Fig. 3.19). We recognized that the
Fourier transform consists of two terms: the section transform result and a
second term resulting from the remainder of the aperture. We derived a
closed-form result for this second term [Eqs. (88) and (89)]. For comparison
to the section results, we plotted the logarithm of the normalized intensity
vs angular coordinate at a fixed radius (Fig. 3.20). The sector pattern
contains two dominant spikes; the third spike is obscured because the
vertical edge of the section aperture was replaced by a circular arc.

Finally, we showed results for the normalized intensity along a radial
line in the transform plane (Figs. 3.21–3.24). We compared three
apertures: a circle, a sector, and a section. We examined the results along
a line corresponding to the obscured sector spike and along a line down the
center of a spike. The section aperture produces spikes whose intensity
drops off monotonically; the sector spikes have a small ripple.

CHAPTER FOUR

DIFFRACTION BY LARGE POLYGONS AND NESTED

POLYGONS

We derive and discuss the diffraction pattern of a nested polygonal aperture in this chapter. The results are new and describe accurately the low-intensity features observed in actual diffraction patterns. A photograph of the aperture for which experimental measurements of the diffraction pattern were made is shown in Fig. 4.1. The transmissive parts of the aperture are pentagon shaped. There are 31 pentagons in the aperture, and the overall size of the actual aperture is 5 mm.

The derivations presented here are based on the result of §3.2.1 [Eq. (3.76)] for a plane-screen corner bounded by a slit aperture. That aperture is an isosceles triangle, and we refer to it as a triangular section of a circular aperture or simply a section. A general polygonal aperture can be broken up into elemental triangles for the purpose of computing its diffraction pattern; if the polygon is regular, then the elemental triangles are sections. Such an analysis has been applied to diffraction by polygonal apertures (Sillitto, 1979), but it is advantageous to review the main points.

The similar decomposition of a nested polygonal aperture into elemental grating-like sections is the key insight that enables us to write a simple closed-form solution for the diffraction pattern of a nested polygonal aperture. First, we derive an expression for the diffraction pattern of a grating-like section and discuss the dominant features. Second, we show how the patterns from grating-like sections are combined to form the
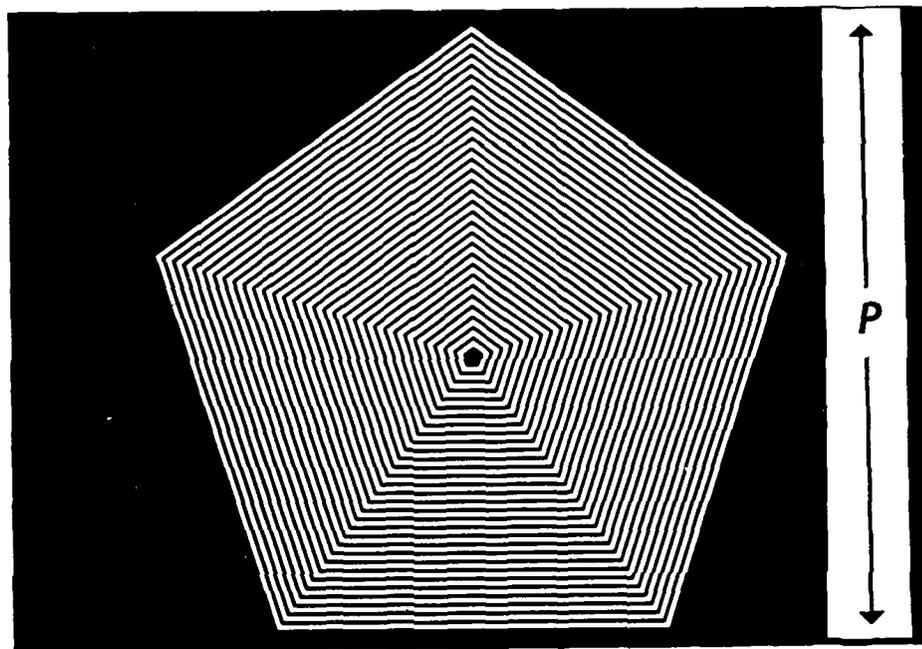
Fig. 4.1. Nested pentagonal aperture used for experimental observations. The transmissive parts of the aperture are white. There are 31 pentagons in the aperture, and the center pentagon is opaque. The overall size of the actual aperture is $P = 5.0$ mm. The other design parameters that characterize the aperture are the altitude or height of a section ($C = 2.24$ mm), the width of the transmissive part ($2a = 0.048$ mm), and the period of the lines ($d = 0.072$ mm).

nested polygon pattern. In order to examine this result in detail, we study a particular case: a nested pentagonal aperture. Computer calculations from our analysis and photographs of actual diffraction patterns reveal low-intensity features and other details. Experimental measurements of the recorded film grain density agree well with the analytic predictions.

4.1 Decomposition of Polygons into Elemental Triangles

In Fig. 4.2 we illustrate the general principle of decomposing the transmission function for an arbitrary aperture, i.e., a polygon, into a sum of transmission functions due to elemental triangles. An interior point of

the polygon, the origin of the xy axes, serves as a common vertex for every triangle. A triangle is traced from the origin to a vertex of the polygon, then to an adjacent vertex, and finally back to the origin. This tracing procedure continues until the entire polygon is described. There will be as many triangles as vertices (or sides) of the polygon. If $f_m^{triangle}(x',y')$ describes the transmission function of the mth triangle and there are n sides, then the transmission function of the polygon $f^{polygon}(x',y')$ is

$$f^{polygon}(x',y') = \sum_{m=0}^{n-1} f_m^{triangle}(x',y') \ . \tag{1}$$

Determining the Fourier transform is straightforward. Taking Fourier transforms on both sides of Eq. (1) gives
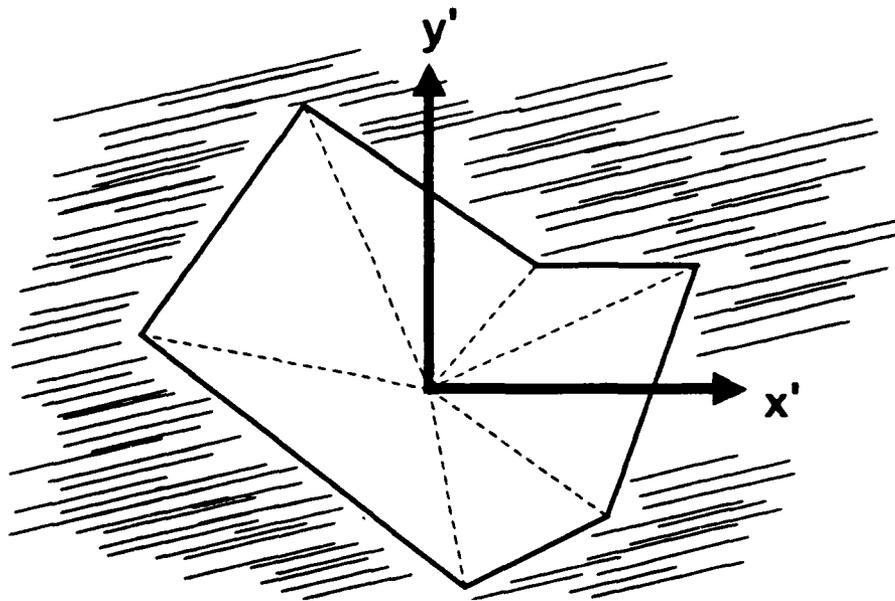
Fig. 4.2. General polygonal aperture. A polygon can be considered as comprised of elemental triangles (dashed lines). An elemental triangle is formed by tracing a closed path from an interior point of the polygon to a vertex, then to an adjacent vertex, and finally back to the interior point. This procedure continues until the entire polygon is described.

$$\bar{f}^{polygon}(u,v) = \sum_{m=0}^{n-1} \bar{f}_m^{triangle}(u,v) \quad , \tag{2}$$

where $\bar{f}_m^{triangle}(u,v)$ can be determined directly. The form of $\bar{f}_m^{triangle}(u,v)$ for a general triangle is not presented here but can be obtained by generalizing the isosceles triangle result from Chapter 3. This technique for polygonal aperture diffraction calculations is due to Sillitto (1979).

An important special case that simplifies the notation occurs when the aperture is an n-sided regular polygon (see Fig. 4.3). In this case all of the elemental triangles are congruent; specifically, they are isosceles triangles with equal opening angles and corresponding sides have equal length. The opening angle θ_0 is given by
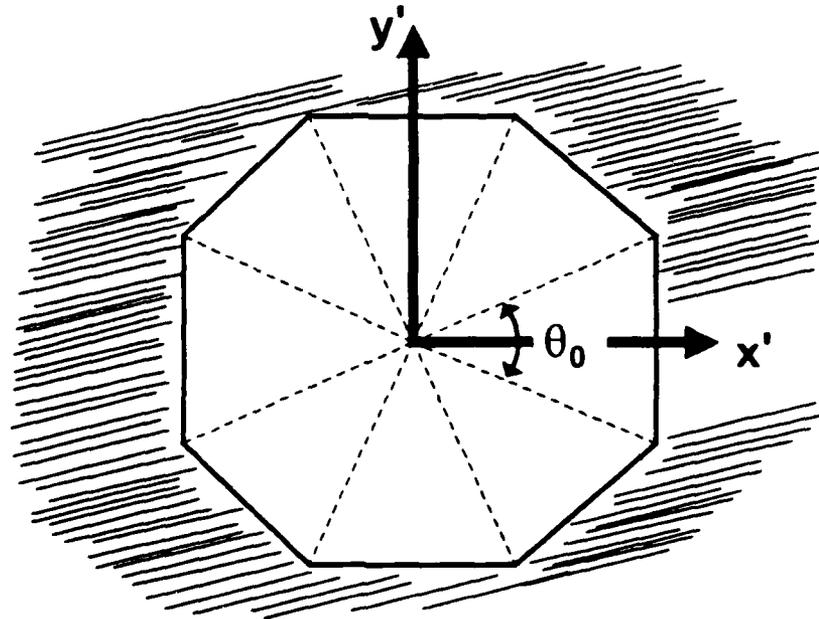
Fig. 4.3. Regular n-sided polygon. In a regular polygon, all of the elemental triangles are isosceles triangles with equal opening angle $\theta_0 = 2n / n = 360° / n$; they are also called sections. Pictured is an octagon with $\theta_0 = 45°$. The sections are indicated by the dashed lines.

$$\theta_0 = 2\pi/n = 360°/n \quad \text{(3)}$$

The elemental triangles are precisely the section apertures studied in §3.2.1. In other words

$$f_m^{triangle}(x',y') = f^{section}(x'_m, y'_m) \quad , \quad \text{(4)}$$

where

$$
\begin{aligned}
x'_m &= r' \cos(\theta' - m\theta_0) \quad , \\
y'_m &= r' \sin(\theta' - m\theta_0) \quad , \\
r' &= (x'^2 + y'^2)^{1/2} \quad , \\
\theta' &= \tan^{-1}(y'/x') \quad ,
\end{aligned}
\quad \text{(5)}
$$

and $f^{section}(x',y')$ is given by Eq. (3.73) with $a = \tan(\theta_0/2)$.

With the transmission function of a regular polygon given by Eqs. (1), (4), and (5), we can write the Fourier transform easily:

$$\bar{f}^{regular\ polygon}(u,v) = \sum_{m=0}^{n-1} \bar{f}^{section}(u_m, v_m) \quad , \quad \text{(6)}$$

$$
\begin{aligned}
u_m &= r \cos(\theta - m\theta_0) \quad , \\
v_m &= r \sin(\theta - m\theta_0) \quad , \\
r &= (u^2 + v^2)^{1/2} \quad , \\
\theta &= \tan^{-1}(v/u) \quad ,
\end{aligned}
\quad \text{(7)}
$$

and $\bar{f}^{section}(u,v)$ is [Eq. (3.76)]

$$\bar{f}^{section}(u,v) = \left(-\frac{C}{i2\pi v}\right)\{\exp(-i2\pi C w_+)\,\text{sinc}(C w_+) - \exp(-i2\pi C w_-)\,\text{sinc}(C w_-)\} \quad , \quad \text{(8)}$$

where

$$w_\pm = u \pm av \quad \text{(9)}$$

To illustrate this decomposition approach to aperture diffraction problems, we present Fig. 4.4. In this figure we display the computed Fourier transform modulus squared or predicted optical transform pattern of an n-sided regular polygon, an octagon, as an 8-bit gray-scale image.

The logarithm of the square modulus of Eq. (6) normalized to unity at the center was computed and mapped to an 8-bit value, i.e., in the range 0 to 255. The mapping is linear in the logarithm such that an intensity of unity yields value 255 (white) and an intensity less than 10^{-7} yields value 0 (b!ack). Hence, we scaled the computed data so that the image represents seven orders of magnitude; this mapping was described in detail in Chapter 3 (see Table 3.2). The parameters used to compute Fig. 4.4 were C = 2.24 mm (the height of a section), θ_0 = 45°, λ = 0.6328 μm, and z_0 = 444 mm. Each pixel corresponds to a 50 μm × 50 μm area in the optical
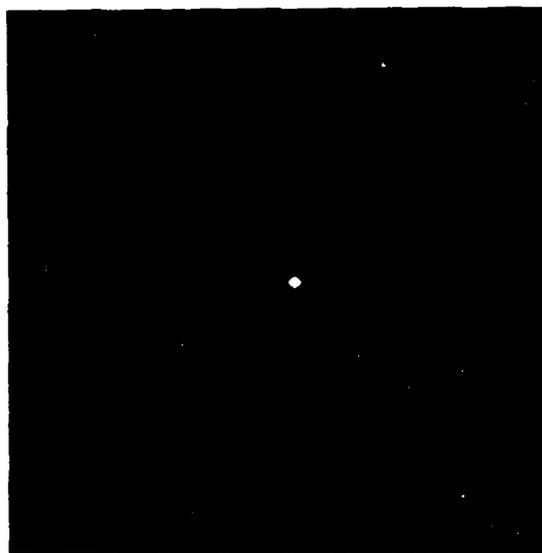
Fig. 4.4. Predicted optical transform pattern of an octagon. The logarithm of the square modulus of Eq. (6) $|f^{octagon}(u,v)|^2$ normalized to unity at the center is displayed as an 8-bit gray-scale image. The mapping of intensity to pixel value is such that an intensity of unity yields pixel value 255 and an intensity below 10^{-7} yields pixel value 0. Details of the intensity mapping (LL7) are presented in the text. The results shown here are for sections with height C = 2.24 mm and θ_0 = 45°. The transform configuration is λ = 0.6328 μm and z_0 = 444 mm. The image is 512 × 512 pixels with each pixel representing an area of 50 μm × 50 μm.

transform plane. Since the image is 512×512 pixels, the observation plane coordinates (x,y) fall in the range (-12.8 mm, 12.8 mm). The frequency variables $(u,v) = (x / \lambda z_0, y / \lambda z_0)$ fall in the range (-45.56 mm^{-1}, 45.56 mm^{-1}). The representation of the diffraction pattern intensity as an 8-bit gray-level image in Fig. 4.4 is similar to the representation in Fig. 3.15. This way of mapping normalized intensity values into pixel values will be referred to as LL7 (linear in the logarithm with range 7).

Clearly, the four intensity spikes dominate the pattern in Fig. 4.4. These are the "light fans" described by Sommerfeld (1954, pp. 233-237). Each edge bounding the octagon produces a spike perpendicular to the edge. Because the number of sides in the polygon is even, the sections are arranged symmetrically about the origin. Hence, each bounding edge can be paired with a parallel edge. The spikes produced by these parallel edges interfere and give rise to the modulation along the spikes. The spatial period of the oscillations is inversely proportional to the separation of the edges ($2C = 4.48$ mm).

It is interesting to note that the spikes produced by edges 1 and 2 of the individual sections (the edges were defined in Fig. 3.14) are not present in the transform pattern. The spike produced by edge 1 of section $m = 0$ is exactly cancelled by the spike produced by edge 2 of section $m = 1$. This is true for the edge 1 and 2 spikes of every section. Only the spikes due to edge 3 of the sections are present in the transform pattern.

4.2   Analysis of Nested Polygons

Now we turn our attention to nested polygonal apertures. In particular, the discussions, derivations, and calculations will be centered on regular polygons, but the generalization to arbitrary polygons is straightforward. In Fig. 4.5 we illustrate the nested polygon to be studied in detail: a pentagon. The dashed lines in (a) denote the sectioning of the aperture. The transform pattern of an individual section (b) will be studied in the following discussion; then this grating-like section pattern will be used to represent the diffraction pattern for the entire nested pentagonal aperture. Finally, we compare experimental measurements obtained from a series of photographs of the diffraction pattern to the computed pattern.

4.2.1   Diffraction by Section of Nested Polygon

The grating-like section shown in Fig. 4.5(b) can be described as a grating function times the section function:

$$f^{gs}(x',y') = f^{grating}(x',y')\, f^{section}(x',y') \quad . \tag{10}$$

The grating function and its transform are discussed below.

The model for the grating is shown in Fig. 4.6. We assume that the grating is infinite along the y axis and periodic along the x axis. The period of the grating lines is d, and the transmissive openings have unity transmission and width $2a$. Further, we assume that the number of grating lines is infinite. The grating is represented as

$$f^{grating}(x',y') = rect\left(\frac{x'}{2a}\right) \otimes_1 \sum_{l=-\infty}^{\infty} \delta(x'-ld) \quad , \tag{11}$$

where $\otimes_1$ denotes a one-dimensional convolution.

123

(a)

(b)
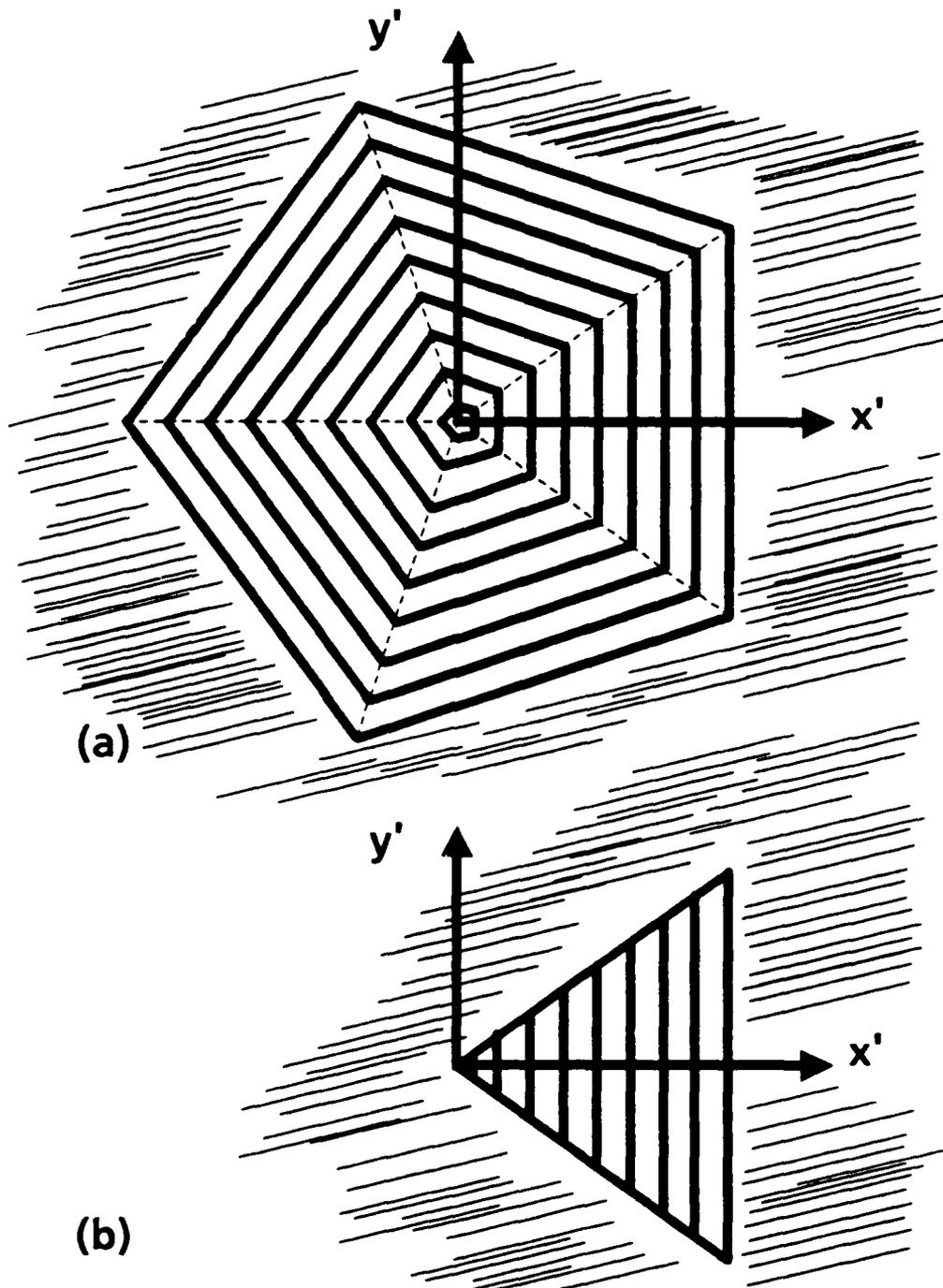
Fig. 4.5. Nested pentagonal aperture studied. In (a) is shown the aperture which consists of nested transmissive pentagons. The dashed lines are not present in the aperture. They are shown to illustrate the elemental building block of the aperture: the grating-like section shown in (b).
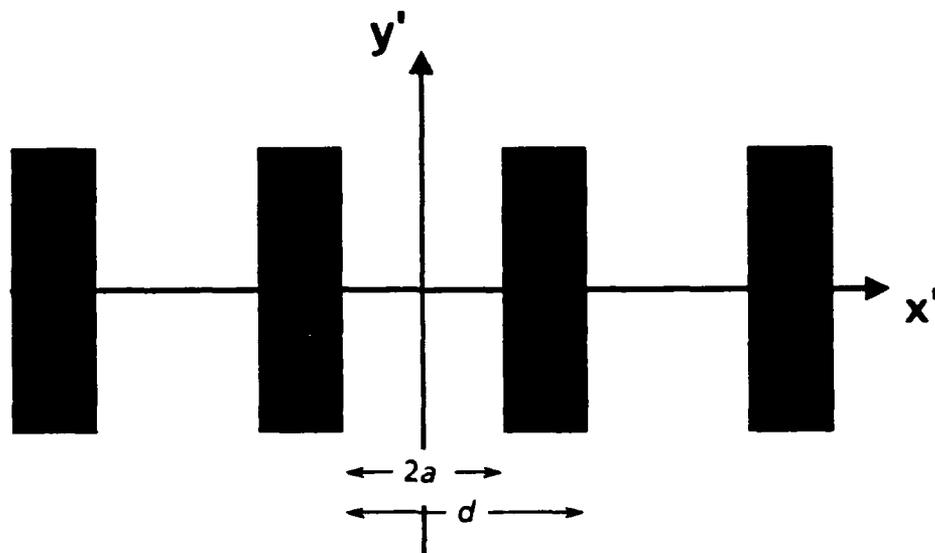
Fig. 4.6. Model for square grating. The grating is infinite in the y direction and has openings of width $2a$ along the x axis. The period of the grating is d. Only a few lines of the grating are shown, but the model assumes that the lines are positioned along the entire x axis.

The Fourier transform can be written in many ways; the following form will be useful for the pending analysis:

$$\bar{f}^{grating}(u,v) = \left(\frac{2a}{d}\right) sinc(2au) \sum_{l=-\infty}^{\infty} \delta\left(u - \frac{l}{d}\right)\delta(v) \quad , \tag{12}$$

where $\delta(u)$ is a one-dimensional Dirac delta function. The characteristics of this transform pattern are well known. The grating produces diffracted orders along the u axis spaced by $1/d$. The strength of these orders is controlled by the ratio $2a/d$. Having an infinite number of lines of infinite length in the grating produces diffracted orders that are two-dimensional Dirac delta functions, i.e., $\delta(u)\delta(v)$.

The point of this analysis is to determine the Fourier transform of Eq. (10); concluding the analysis requires straightforward application of the

sifting property of delta functions to the two-dimensional convolution integral that results from the Fourier transform of a product of two functions. In other words, taking the Fourier transform of Eq. (10) involves the following steps.

$$\bar{f}^{gs}(u,v) = \bar{f}^{grating}(u,v)\otimes_2 \bar{f}^{section}(u,v) \qquad ,$$

$$\bar{f}^{gs}(u,v) = \left(\frac{2a}{d}\right)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} sinc[2a(u-u')]$$

$$\times \sum_{l=-\infty}^{\infty} \delta\left[(u-u')-\frac{l}{d}\right]\delta(v-v')\bar{f}^{section}(u',v')\,du'\,dv' \quad ,$$

$$\bar{f}^{gs}(u,v) = \left(\frac{2a}{d}\right)\sum_{l=-\infty}^{\infty} sinc\left(\frac{2al}{d}\right)\bar{f}^{section}\left(u-\frac{l}{d},v\right) \quad . \tag{13}$$

Equation (13) is the Fourier transform of the grating-like section of Fig. 4.5(b). The structure of the equation reveals that the transform of the section is replicated at the position of each of the diffracted orders of the grating. The contributions from each order are added coherently with a weighting governed by the ratio $2a/d$.

Figure 4.7 is an LL7 representation of the transform pattern of the grating-like section. The parameter values used to compute the figure were C = 2.24 mm, θ_0 = 72°, d = 0.072 mm, $2a$ = 0.048 mm, λ = 0.6328 μm, and z_0 = 444 mm. At each order of the grating is centered the section transform pattern (albeit modified somewhat by the interference from other orders). Interestingly, the third diffracted order is missing. This is due to the value of the ratio $2a/d$ = 2/3. At every third order (l = ±3, 6, 9, ···) the *sinc* factor in the summation is zero; hence, there is no
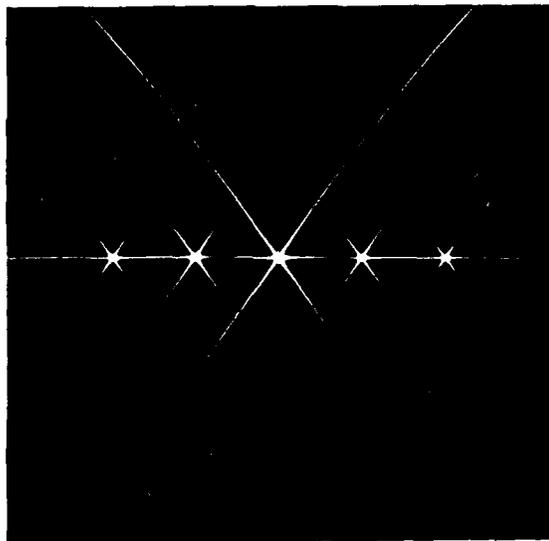
Fig. 4.7. Optical transform pattern of grating-like section. The logarithm of the square modulus of Eq. (13) $|f^{gs}(u,v)|^2$ normalized to unity at the center is displayed as an 8-bit gray-scale image (LL7 representation). The result shown here is for a section with height $C = 2.24$ mm and $\theta_0 = 72°$. The grating parameters are $d = 0.072$ mm and $2a = 0.048$ mm. The transform configuration is $\lambda = 0.6328$ µm and $z_0 = 444$ mm.

contribution to the transform from these orders. Further, the figure shows

that there is significant diffracted intensity in regions away from the

center. In other words, the diffraction features produced by the

combination of the grating and the triangular shape are not restricted to

the central region of the pattern. This effect contributes greatly to the

nested polygon diffraction pattern.

4.2.2   Nested Pentagons: Computer Calculations

All of the subsidiary analytic results are in hand to perform the

desired calculation of the Fourier transform of a nested pentagonal

aperture, such as the one shown in Fig. 4.1. Such a computation relates directly to predicting the observed intensity distribution in an optical transform configuration. The construction of the solution is analogous to the construction of the solution for a simple polygon, Eqs. (6) and (7): the elemental grating-like sections are summed. The solution for the nested pentagon is

$$\bar{f}^{nested\ pentagon}(u,v) = \sum_{m=0}^{4} \bar{f}^{gs}(u_m,v_m) \ ,\tag{14}$$

where

$$
\begin{aligned}
u_m &= r\cos(\theta - m\theta_0) \ , \\
v_m &= r\sin(\theta - m\theta_0) \ , \\
r &= (u^2+v^2)^{1/2} \ , \\
\theta &= \tan^{-1}(v/u) \ ,
\end{aligned}\tag{15}
$$

and

$$\theta_0 = 360°/5 = 72° \ .\tag{16}$$

$\bar{f}^{gs}(u,v)$ is given by Eq. (13), and $\bar{f}^{section}(u,v)$ is given by Eqs.(8) and (9) with

$$a = \tan(\theta_0/2) = 0.7265 \ .\tag{17}$$

From a computational standpoint the solution is especially straightforward. First one creates a module that computes $\bar{f}^{section}(u,v)$ according to Eqs. (8) and (9). A second module computes $\bar{f}^{gs}(u,v)$ according to Eq. (13) using the first module. A final module computes the transform of the nested pentagonal aperture acording to Eqs. (14) and (15) by means of the other two modules.

We show an LL7 representation of the computed transform pattern in Fig. 4.8. The aperture and transform parameters are $C = 2.24$ mm, $d = 0.072$ mm, $2a = 0.048$ mm, $\lambda = 0.6328$ µm, and $z_0 = 444$ mm. There are 31 transmissive pentagons in the aperture and the overall size of the aperture

is 5.0 mm. These parameters were the design parameters specified for manufacturing the nested pentagonal aperture shown in Fig. 4.1.

Examination of Fig. 4.8 reveals the presence of some rather striking features. First, the five grating-like sections have produced rings of diffracted orders; there are ten orders in a ring. Emanating from each order are three spikes. (These spikes are the replication of the section transform pattern at each diffracted order.) The radial spikes in the overall pattern are attributable to the bounding edges of the pentagon (referred to as edge 3 in the discussion of the section pattern). The other two spikes (due to edges 1 and 2) trace out pentagons! At each ring there
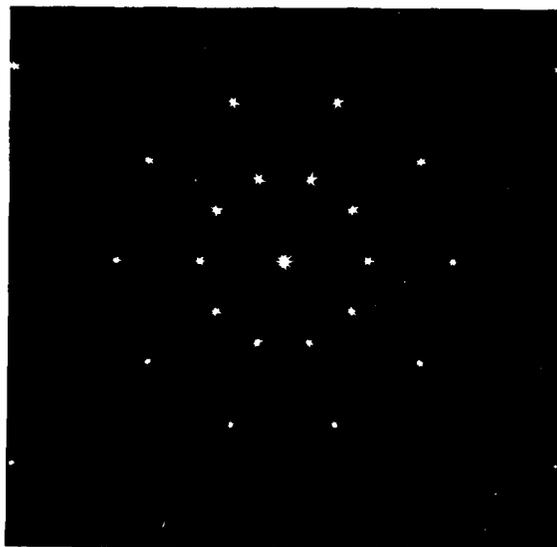
Fig. 4.8. Predicted optical transform pattern of nested pentagonal aperture computed from Eq. (14). The logarithm of the square modulus of Eq. (14) $|f^{nested\,pentagon}(u,v)|^2$ normalized to unity at the center is displayed as an 8-bit gray-scale image (LL7 representation). The result shown here is for a section with height C = 2.24 mm and θ_0 = 72°. The grating parameters are d = 0.072 mm and $2a$ = 0.048 mm. The transform configuration is λ = 0.6328 μm and z_0 = 444 mm.

are two discernible pentagons. In the following discussion we elucidate and explain this feature.

Suppose that instead of the square grating described by Eq. (11), the grating is cosinusoidal, i.e.,

$$f^{cosine\ grating}(x',y') = \cos\left(\frac{2\pi x'}{d}\right) \quad . \tag{18}$$

Then, the grating transform has only two orders, i.e.,

$$\bar{f}^{cosine\ grating}(u,v) = \frac{1}{2}\left[\delta\left(u-\frac{1}{d}\right) + \delta\left(u+\frac{1}{d}\right)\right]\delta(v) \quad . \tag{19}$$

Retracing the steps leading up to Eq. (13) gives, for a cosine grating-like section,

$$\bar{f}^{cgs}(u,v) = \frac{1}{2}\left[\bar{f}^{section}\left(u-\frac{1}{d},v\right) + \bar{f}^{section}\left(u+\frac{1}{d},v\right)\right]\delta(v) \quad . \tag{20}$$

To construct a simplified picture of the modulus of $\bar{f}^{cgs}(u,v)$, we represent $\bar{f}^{section}(u,v)$ by three lines drawn at angles corresponding to the spikes in the transform pattern. We represent Eq. (20) by drawing two sets of spikes at $u = \pm 1/d$. Fig. 4.9(a) shows this. (Only the spikes due to edges 1 and 2 are drawn. We omit the third spike for clarity in the figure.) We show the set of spikes at $u = +1/d$ with dashed lines and the set of spikes at $u = -1/d$ with solid lines.

A cosine grating-like section rotated by 72° produces the same transform pattern as (a) but rotated 72° [see Fig. 4.9(b)]. Each section of the cosine nested aperture contributes similarly [see Figs. 4.9(c)–(e)]. The sum of these five patterns is shown in Fig. 4.9(f). Although we obtained (f) by adding incoherently the five section patterns, we can see that some features of the computed diffracted pattern in Fig. 4.8 are displayed.
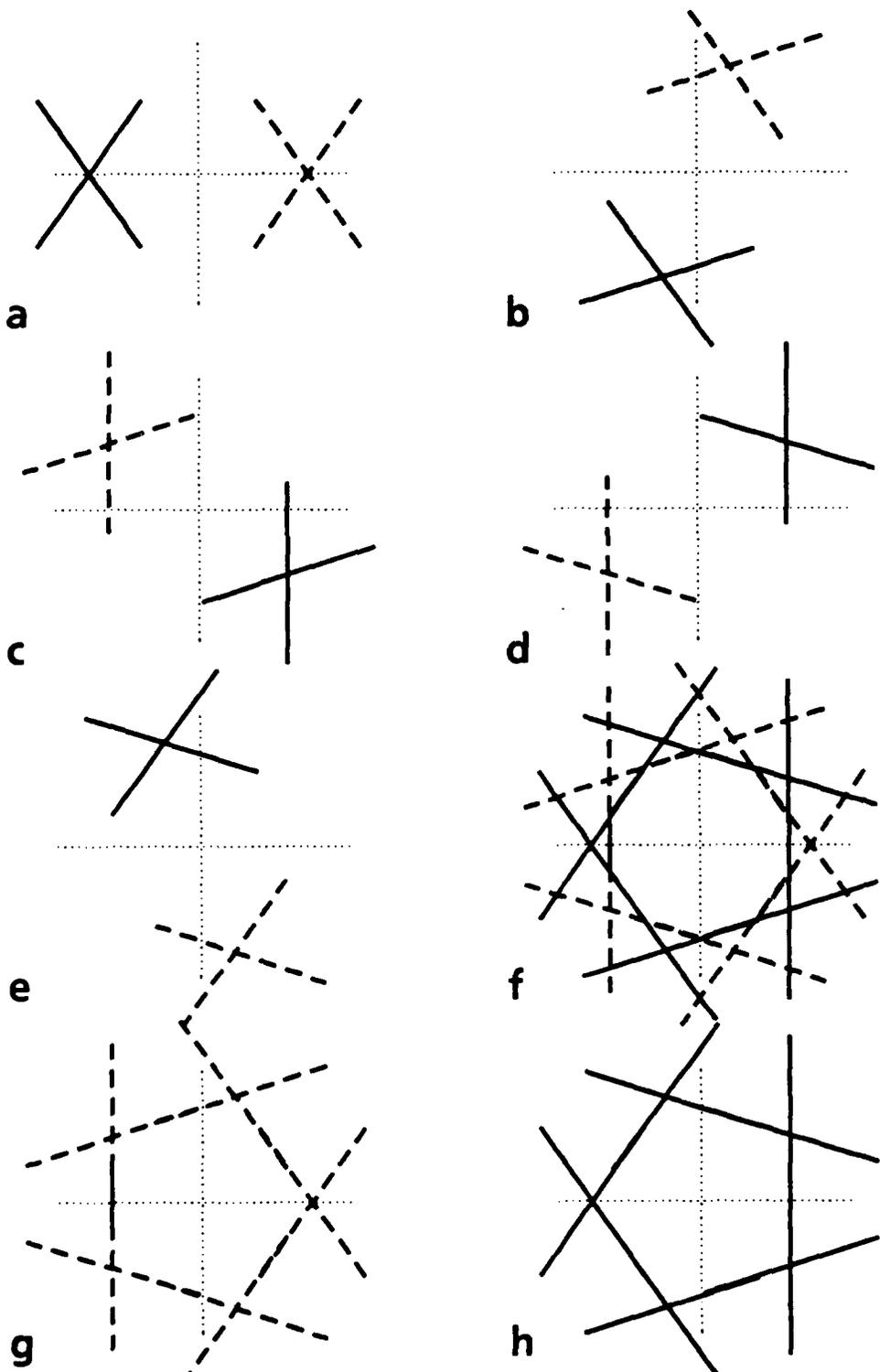
Fig. 4.9. Graphical representation of cosine grating-like section transform patterns used to explain nested pentagon transform features.

The reason for representing the section pattern at $u = +1/d$ with dashed lines is apparent, especially considering Figs. 4.9(g) and (h). The contribution from the right-hand side set of spikes yields a pentagon; the left-hand set of spikes yields a pentagon reflected about the v axis. A similar graphical construction can be conducted at each ring of diffracted orders to discover that the transform of a nested pentagonal aperture contains double nested pentagons traced out by the intensity spikes.

Of course, the graphical construction ignores interference between the orders and says nothing about the relative magnitude of the intensities along the spikes, but it provides a simple way to understand what parts of the aperture transmission function produce identifiable features in the transform pattern. The precise solution for the transform pattern, Eq. (14) and related equations, predicts the relative intensities and includes interference effects.

To investigate the relative magnitude of the intensity along the spikes and to investigate how the diffraction pattern depends on the aperture parameters, we consider single line scans through the predicted or computer-generated diffraction pattern. In other words, we sample the transform pattern along a line of interest. Consider a horizontal scan through the pattern at $y = 0$ mm (–12.775 mm $\leq x \leq$ 12.775 mm). Figure 4.10 is a duplication of Fig. 4.8 except that the lines to be sampled are indicated. This first scan line of interest connects points P1 and P2.

In Fig. 4.11 we show a horizontal scan through the predicted transform pattern for two values of d, the grating spacing. The logarithm of the normalized intensity along the P1–P2 line for the pattern shown in
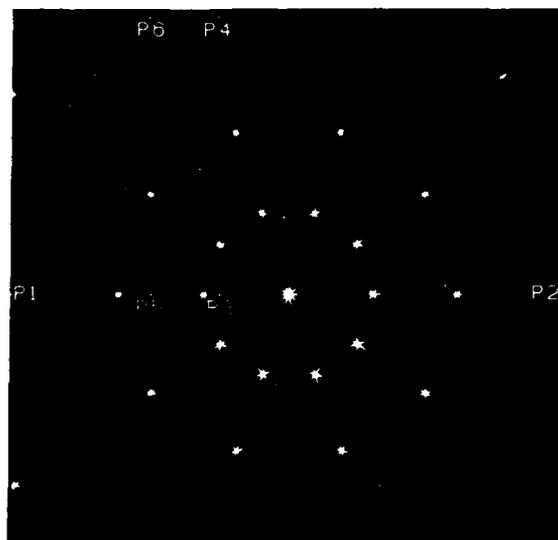
Fig. 4.10. Predicted optical transform pattern of nested pentagonal aperture showing lines along which the pattern is sampled. The figure is a duplication of Fig. 4.8. Three scans are examined: P1–P2, P3–P4, and P5–P6.

Fig. 4.8, i.e., d = 0.072 mm, is contrasted with the result for a larger value of d (= 0.075 mm). In this second curve we see how the diffracted orders shift towards the central peak as d increases. The intensity between the first orders is about 50 dB down from the central peak intensity for both cases. In the calculations, the ratio $2a/d$ remained constant ($2a/d$ = 2/3).

In Fig. 4.12 we see the effect of varying the ratio $2a/d$. Both curves are for d = 0.072 mm, but in one case $2a$ = 0.048 mm and in the other case $2a$ = 0.047 mm. The ratio controls the relative magnitude of the diffracted orders. The first two orders are very nearly equal for the different ratios, but varying $2a/d$ even slightly from 2/3 produces a third diffracted order.
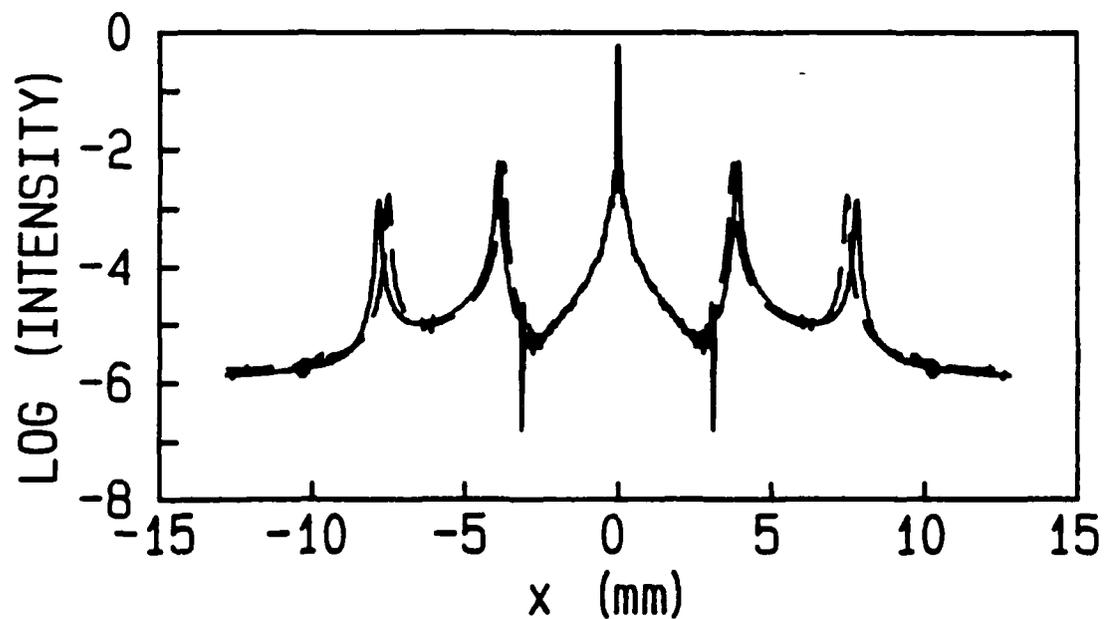
Fig. 4.11. P1–P2 scan through the predicted nested pentagon diffraction pattern for $d = 0.072$ mm (——) and $d = 0.075$ mm (- - - -). The ratio $2a/d$ is the same for both curves ($= 2/3$).
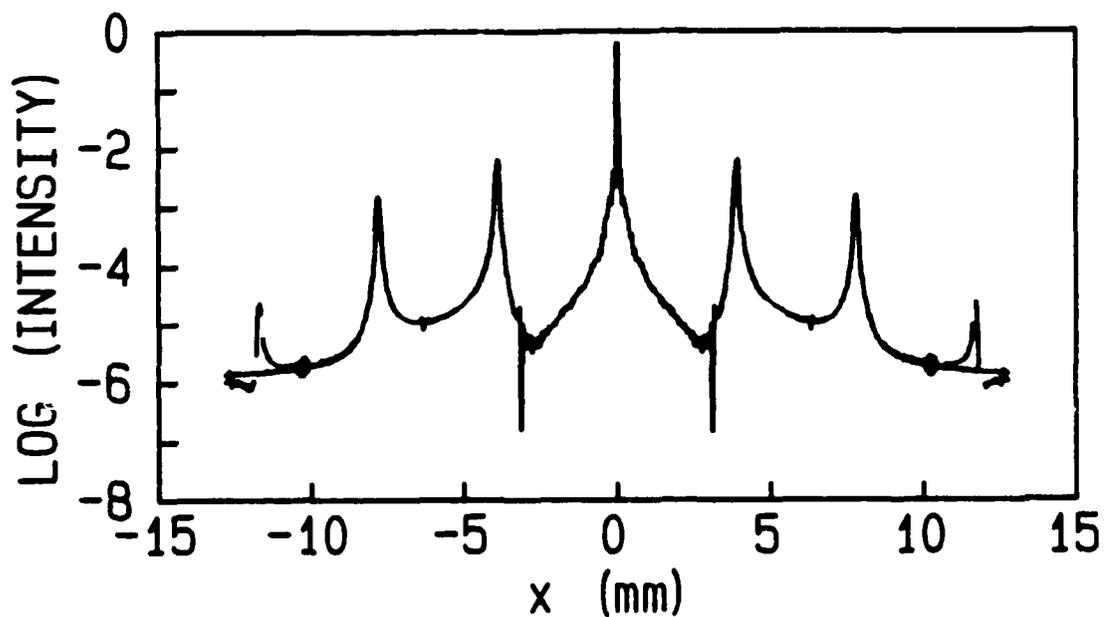
Fig. 4.12. P1–P2 scan through the predicted nested pentagon diffraction pattern for $2a = 0.048$ mm (——) and $2a = 0.047$ mm (- - - -). The grating spacing d is the same for both curves ($= 0.072$ mm).

The height of a nested section, C, influences the pattern as well. Figure 4.13 contains the design criteria plot (C = 2.24 mm) and a plot for C = 2.254 mm. The diffracted orders are not affected by this variation, but the region between the orders is. At the larger value of C, the intensity level between the zero and first order increases to less than 50 dB down, but the intensity level between the first and second order decreases to about 70 dB down. The influence of the value of C on the intensity at one point, e.g., (x, y) = (–6.525 mm,0 mm), in the diffraction pattern is seen in Fig. 4.14. The logarithm of the normalized intensity is plotted vs C for 2.14 mm $\leq C \leq$ 2.34 mm. The plot shows that the intensity at a single point can vary over two orders of magnitude with a change in C of 0.02 mm.

It is also interesting to see how the pattern changes with a small shift in observation coordinates. In Fig. 4.15 we compare the result of moving the scan line up to y = 0.05 mm to the original scan line plot. The curve is shifted down approximately 5 dB, but retains the same general features.

It is clear that the details of the predicted intensity along this scan line are very sensitive to small changes in the aperture parameters and observation coordinates. Calculations through lines not including the center of the pattern are not as sensitive. A scan following P3–P4 in Fig. 4.10 (x = –3.20 mm, 0 mm $\leq y \leq$ 12.775 mm) is shown in Fig. 4.16. The plot has a maximum at the location of a first diffracted order; subsidiary maxima are encountered as different spikes are crossed. Figure 4.17 shows a scan along P5–P6 (x = –6.35 mm, 0 mm $\leq y \leq$ 12.775 mm). The major peak of the scan occurs at a second diffracted order; the crossing of other
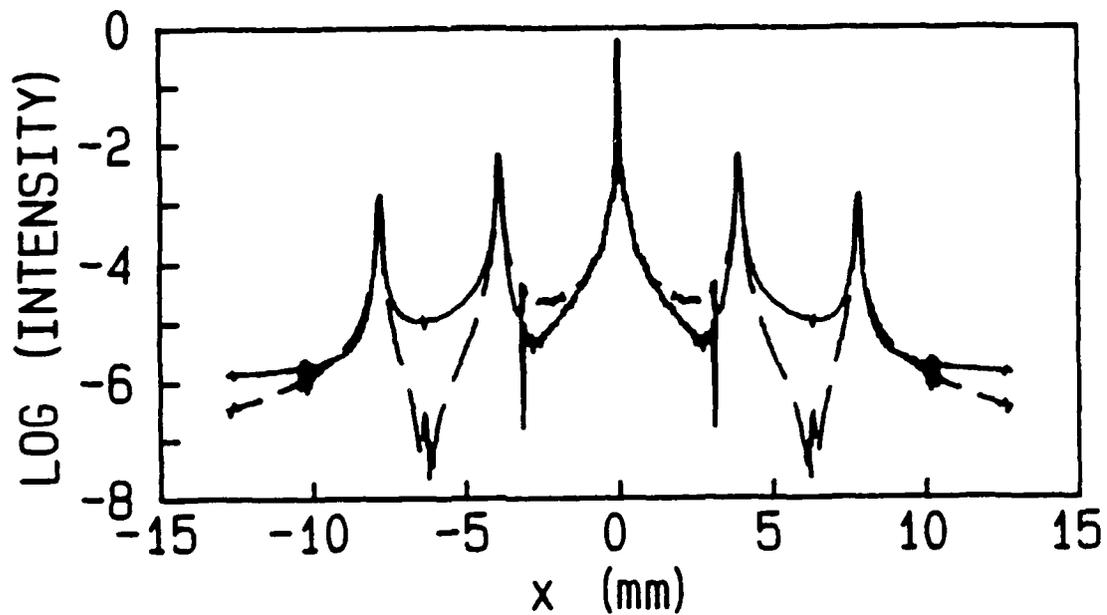
Fig. 4.13. P1–P2 scan through the predicted nested pentagon diffraction pattern for $C = 2.24$ mm (———) and $C = 2.254$ mm (- - - -).
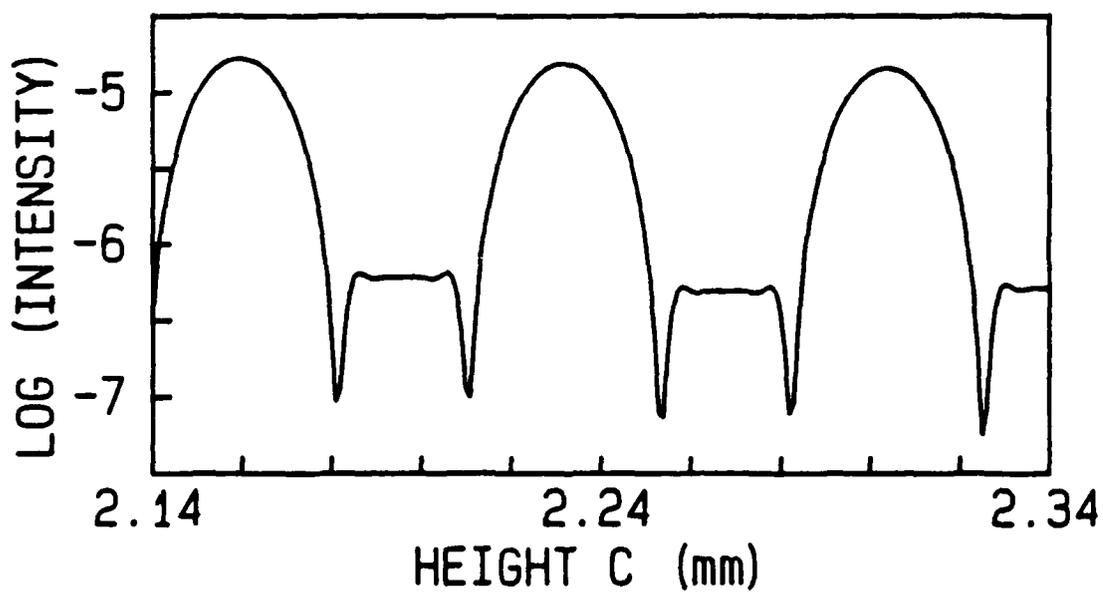
Fig. 4.14. Logarithm of normalized intensity at a single point, $(x,y) = (-6.525$ mm,0 mm), vs the height of the section C.
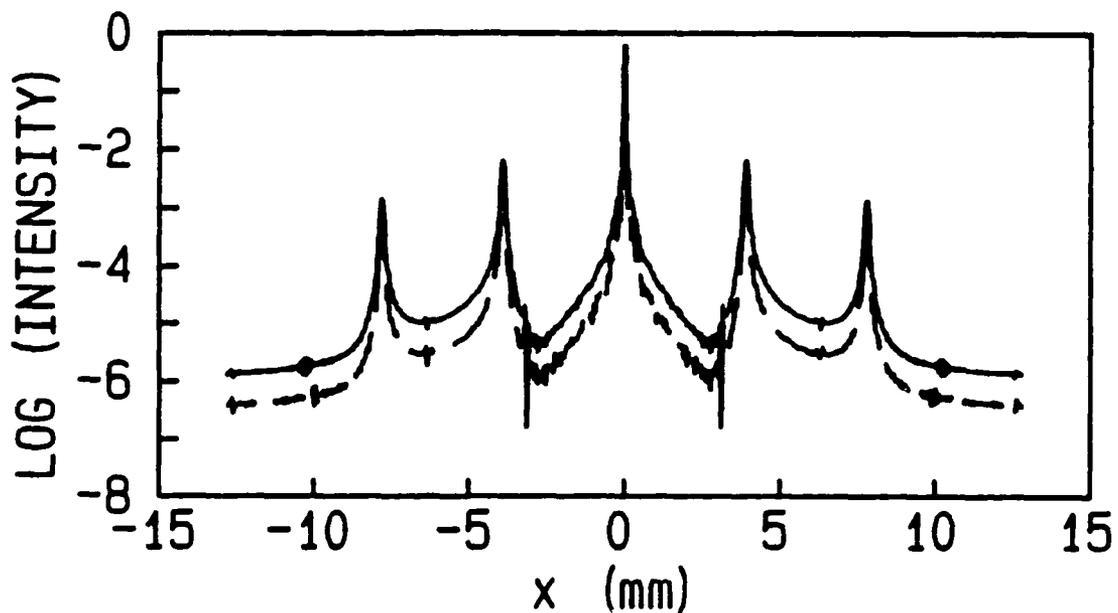
Fig. 4.15. P1–P2 scan through the predicted nested pentagon diffraction pattern (———) contrasted with a similar scan with $y = 0.05$ mm (- - - -).
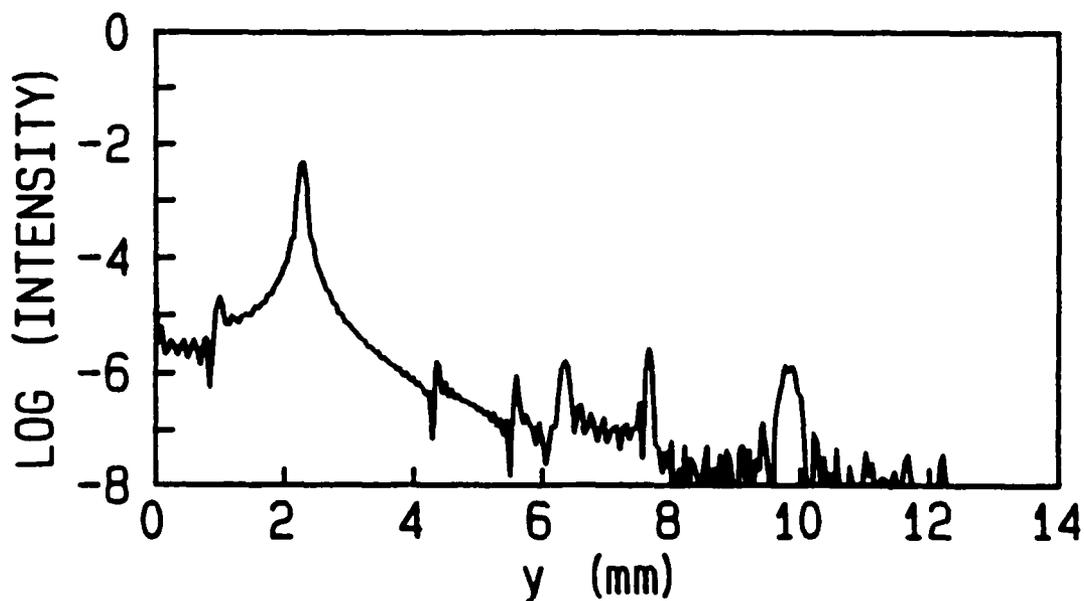
Fig. 4.16. P3–P4 scan through the predicted nested pentagon diffraction pattern for $C = 2.1814$ mm, $d = 0.072$ mm, and $2a = 0.0475$ mm.
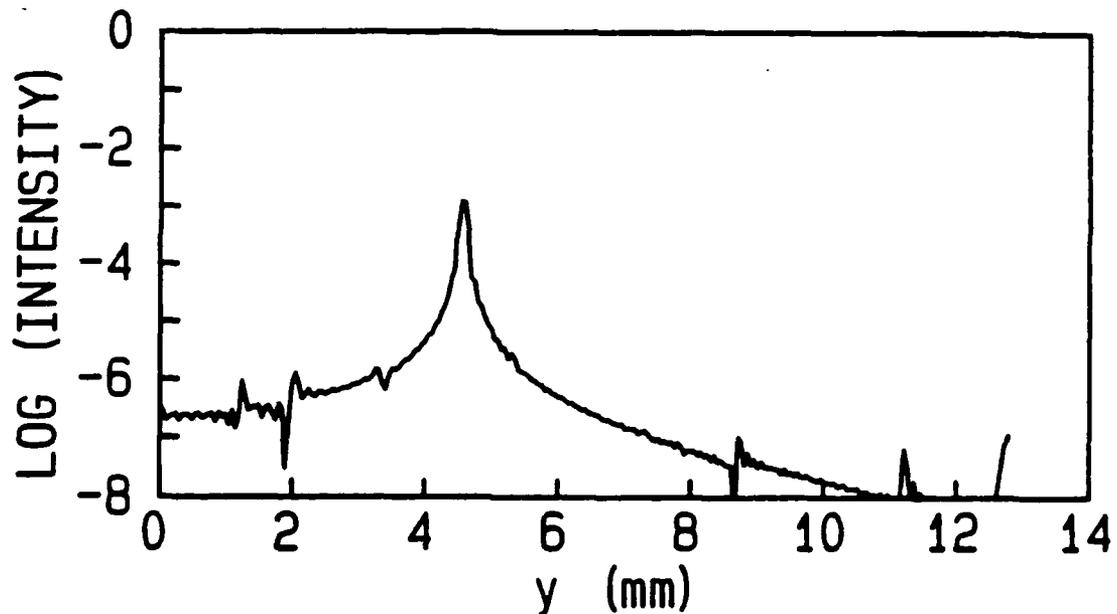
Fig. 4.17. P5–P6 scan through the predicted nested pentagon diffraction pattern for $C = 2.1814$ mm, $d = 0.072$ mm, and $2a = 0.0475$ mm.

spike lines is not as noticeable because of the decreased overall intensity level. We computed the two preceding figures for an aperture in a transform system with the following parameters: $C = 2.1814$ mm, $d = 0.072$ mm, $2a = 0.0475$ mm, $\lambda = 0.6328$ µm, and $z_0 = 444$ mm.

4.2.3   Nested Pentagons: Photographs and Diffraction Features

As a first step in confirming the validity of the computational results, we photographed the diffraction pattern of a nested pentagonal aperture in a converging, monochromatic beam. The setup used to record the optical transform pattern is shown in Fig. 4.18. A HeNe laser (L) was attenuated by neutral density filters (NDF). The beam was expanded from ~3 mm diameter to ~25 mm diameter by the spatial filter (SF) – collimating lens
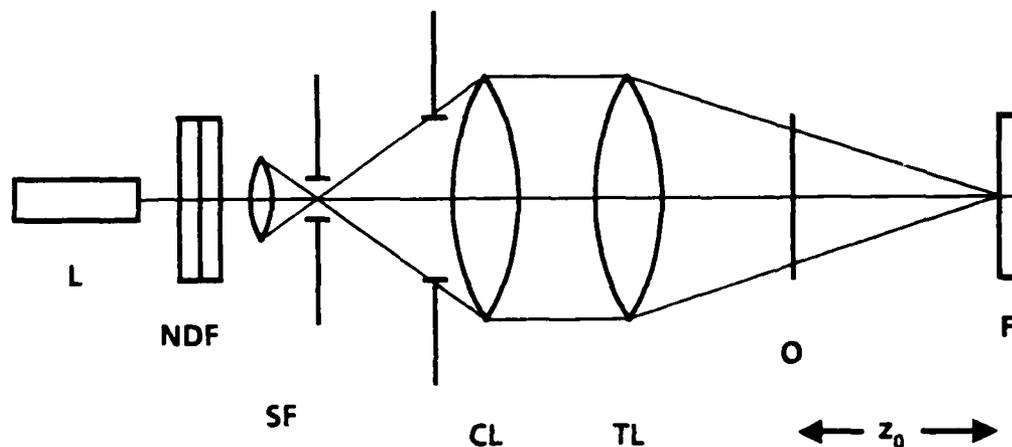
Fig. 4.18. Experimental arrangement for photographically recording diffraction patterns. Monochromatic light from the laser (L) was attenuated by the neutral density filters (NDF) and was expanded by the spatial filter (SF) – collimating lens (CL) combination. The transform lens (TL) created a converging spherical beam of unit amplitude that illuminated the object (O) and formed the transform or diffraction pattern on the film (F). The distance z_0 determined the scaling of the transform.

(CL) combination. The collimated beam was brought to a focus by the

transform lens (TL). The converging beam illuminated the object (O) and

the transform or diffraction pattern was recorded on film (F). The HeNe

laser used was a Spectra-Physics model 102-3; the output was ~5 mW,

linearly polarized with $\lambda = 0.6328\ \mu m$. The neutral density filters were

2″ × 2″ glass slides; the optical density of a slide was either 0.3 or 1.0;

several slides were stacked if necessary to achieve the desired attenuation.

The spatial filter was a standard microscope objective (40×, NA = 0.65) –

pinhole (10 μm) arrangement. The collimating lens was a plano-convex

lens with a focal length of ~500 mm. The plano-convex lens is a typical lens

used in collimators and beam **expanders** because spherical aberration is minimized for a singlet at infinite conjugate ratio. To ensure uniform amplitude of the collimated beam produced, we centered an aperture of diameter ~25 mm on the collimating lens. Hence, only the central portion of the greatly overexpanded laser beam was used. Collimation of the beam was checked by means of a lateral shearing plate.

A long focal length transform lens (f $\approx$ 750 mm) allowed wide latitude in selecting the scale of the optical transform, i.e., z_0. The transform lens was a cemented telescope doublet. This choice was made because such a doublet is well corrected for spherical aberration. If the optical system is properly aligned, this is the predominant aberration. (The optics were carefully aligned by examining the back reflections from their surfaces.) Hence, the converging wave illuminating the object was spherical to high accuracy.

Various transform configurations and the performance of various types of lenses as transform elements have been discussed (Kane, 1984; Joyeux and Lowenthal, 1982; Casasent and Luu, 1978a, b). In the most general case Joyeux and Lowenthal showed that the converging beam transform (see Fig. 4.10) is preferable when off-the-shelf lenses are used.

A Nikon F3 camera body with data back and motor drive was positioned in the focal plane. This camera featured accurate metering with the lens removed and shutter speeds from 1/2000 second to 8 seconds, stepping by a factor of two. Thus, photographs over a wide range of exposures or dynamic range were possible. The motor drive permitted successive frames to be recorded with little or no intermittent camera

movement, and the data back imprinted a date stamp useful for image registration. The standard technique of placing an edge in the object plane to locate the focal plane was used. KODAK T-MAX 100 film was loaded into camera. The development of the film is discusssed in detail in Appendix D, but it was such that intensities over ~3 orders of magnitude could be recorded. This film characteristic plus the dynamic range of the shutter speeds enabled us to observe and measure intensities over ~7 orders of magnitude.

The nested pentagonal aperture used was a $2'' \times 2''$ chrome on glass mask (see Fig. 4.1). The specifications for fabrication were overall size $P =$ 5.0 mm (hence, $C = 2.24$ mm), $d = 0.072$ mm, and $2a = 0.048$ mm. Measurements of the aperture under a microscope yielded the following results: $C = 2.17 \pm .02$ mm, $d = 0.070 \pm .002$ mm, and $2a = 0.045 \pm .002$ mm. The optical density of the chrome was ~3.0, and the chrome side of the mask was oriented toward the illuminating beam.

Neutral density filters were inserted so that a somewhat overexposed pattern on the film was obtained for ~1/8 second shutter speed; a neutral density of 2.0 achieved this. Thus, the bracketing of exposures from 1/2000 to 8 seconds provided a useful range of observation. It was desired that no saturation of the film occur at the shortest shutter exposure time (1/2000 second achieved this) and that stray light be minimal at the longest exposure time. The 8 second exposure displayed noticeable stray light; hence, it was not used. The 4 second exposure was the longest exposure used.

Figure 4.19 shows a sample of the photographs taken. Photograph (a)
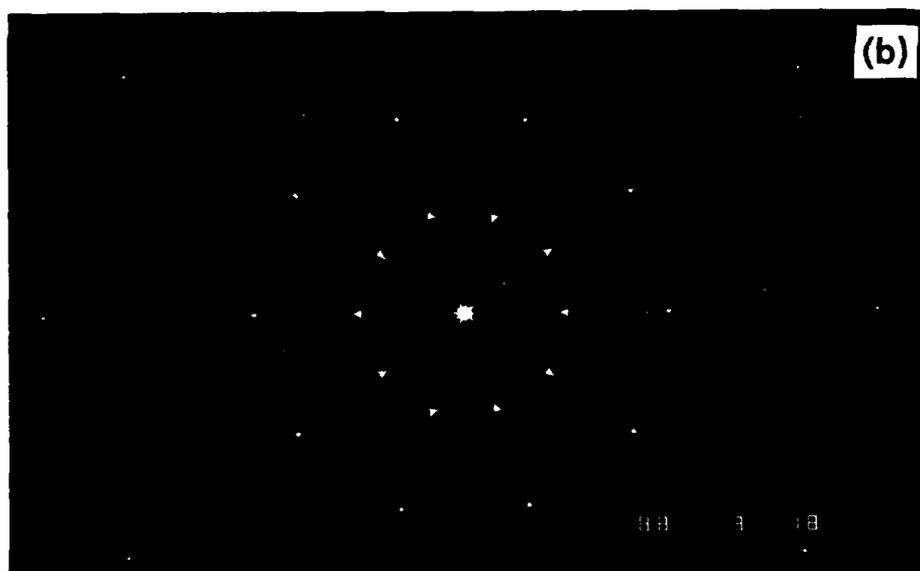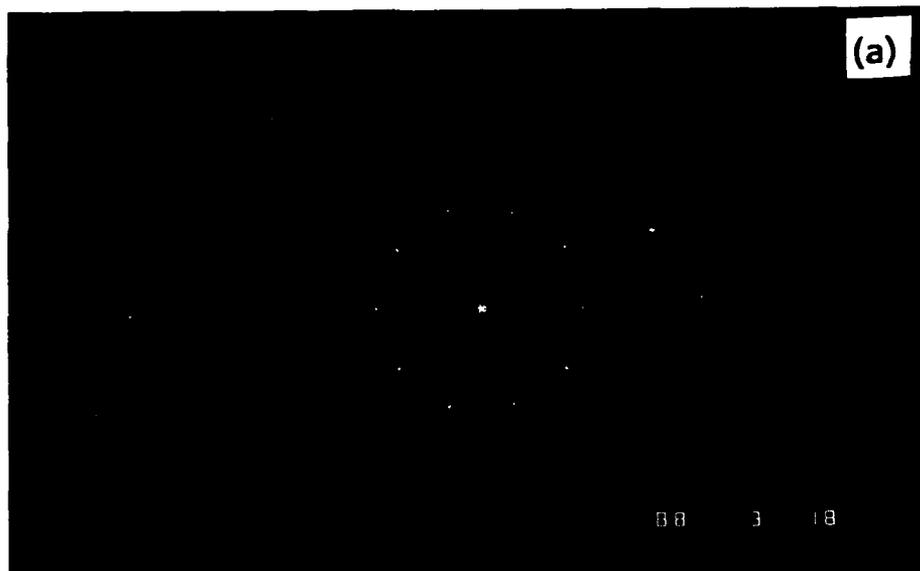
Fig. 4.19. Experimental photographs of optical transform of nested pentagonal aperture. The aperture of Fig. 4.1 is transformed and the intensity pattern is recorded on film. Shown are various exposure times: (a) 1/1000 second and (b) 1/60 second. The scaling distance is $z_0 = 444$ mm. See text for discussion of diffraction pattern features.
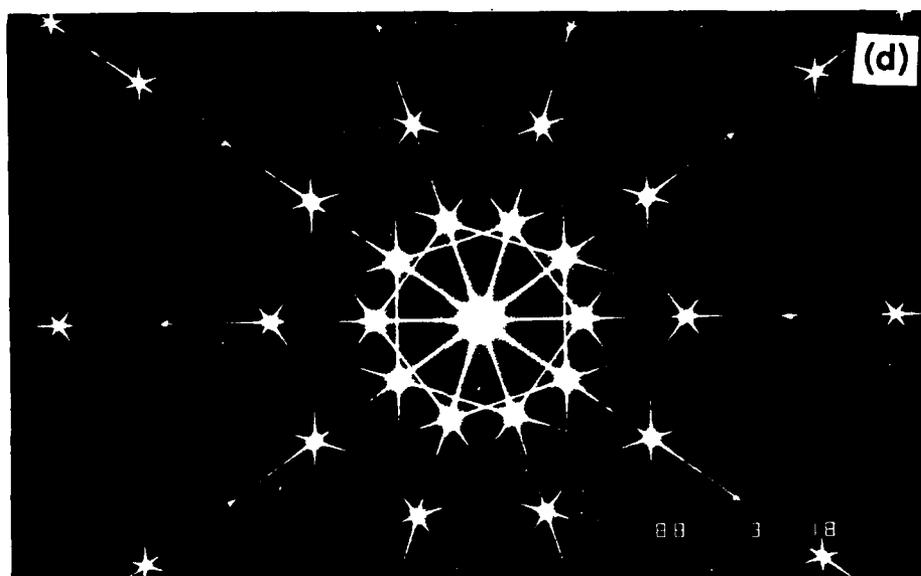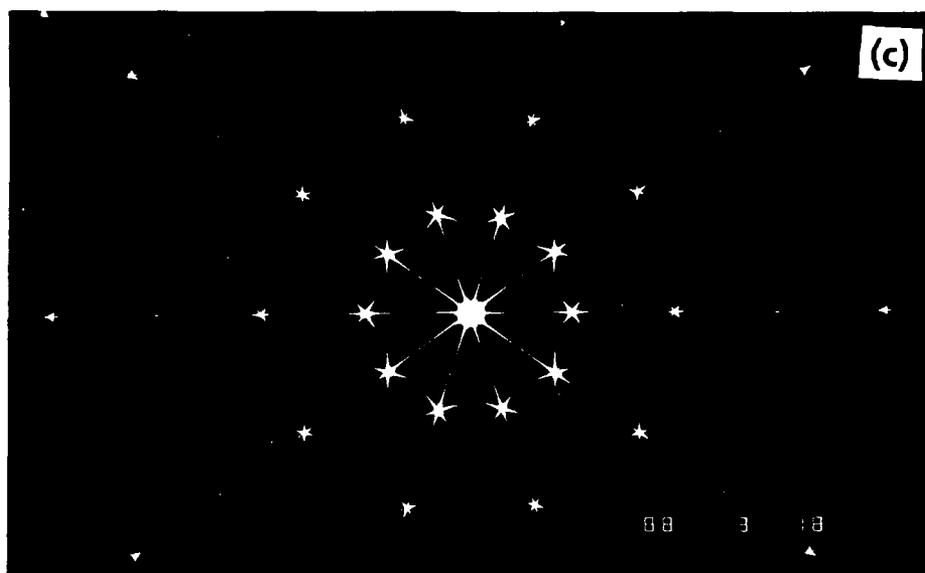
142

Fig. 4.19. Experimental photographs of optical transform of nested pentagonal aperture. The aperture of Fig. 4.1 is transformed and the intensity pattern is recorded on film. Shown are various exposure times: (c) 1/4 second and (d) 4 seconds. The scaling distance is $z_0 = 444$ mm. See text for discussion of diffraction pattern features.

is 1/1000 second, (b) is 1/60 second, (c) is 1/4 second, and (d) is 4 seconds. The numbers 88 3 18 in the lower right-hand corner are the date stamp supplied by the data back. We used them for image registration and reference in the quantitative analysis to follow.

At the shortest exposure, we can discern only the diffracted orders; none of the intensity spikes connecting the orders is yet manifest. These spikes become apparent with increasing exposure time and are especially pronounced in the 4 second exposure. We can trace out the pentagons in the transform pattern at the first and second diffracted orders. The third diffracted order is visible, but noticeably lower in intensity. The ratio $2a/d$ is only approximately equal to 2/3. We chose $z_0 = 444$ mm so that ± 3 orders fit vertically in the 35 mm frame. This scale was advantageous in performing the detailed analysis of the features of the diffraction pattern in the following section.

From a qualitative standpoint, the photographs in Fig. 4.19 confirm the existence of the features discussed in §4.2.2, Fig. 4.8. The transform pattern displays rings of diffracted orders, each order having three spikes. At high exposure times, the intensity spikes trace out the nested pentagon shapes explained in Fig. 4.9. The quantitative agreement is investigated in the following section.

4.2.4   Nested Pentagons: Experimental Measurements and Comparison to Calculations

The data to be analyzed were recorded in raw form as the film grain density in each of the exposed frames. We used photographic film to record

the transform data because of its high spatial resolution and wide dynamic range over a two-dimensional sampling space. These attributes were especially advantageous in view of the characteristics of the pattern to be analyzed.

Determining quantitatively how film grain density relates to intensity was one of the difficulties in comparing the experimental data to the theoretical calculations. A sensible model for this determination is the Hurter-Driffield curve or D-log-e curve of the film; such a curve quantifies and relates optical density of the film to the logarithm of exposure. Here, exposure, e, means

$$e = EE^* t \quad , \tag{21}$$

where EE^* is the intensity of the field (the square modulus of the electric field), and t is the duration of the exposure. The details are explained in Appendix D, but the main point is that measurement of the film grain density over an area large compared with the size of a film grain, but small compared with an area within which the density changes significantly, combined with a plot of optical density vs the logarithm of exposure indicated the local average of the intensity over that area.

A block diagram of the experiment is displayed in Fig. 4.20. The intensity (EE^*) in the transform plane was recorded on film for an exposure time t. After development, the film was scanned and digitized on a Photomation P-1700 (Optronics International, Inc.) high-speed digital scanner. The scanned information was transferred to the host computer via computer tape. We performed various software manipulations using a commercial image processing system (International Imaging Systems,
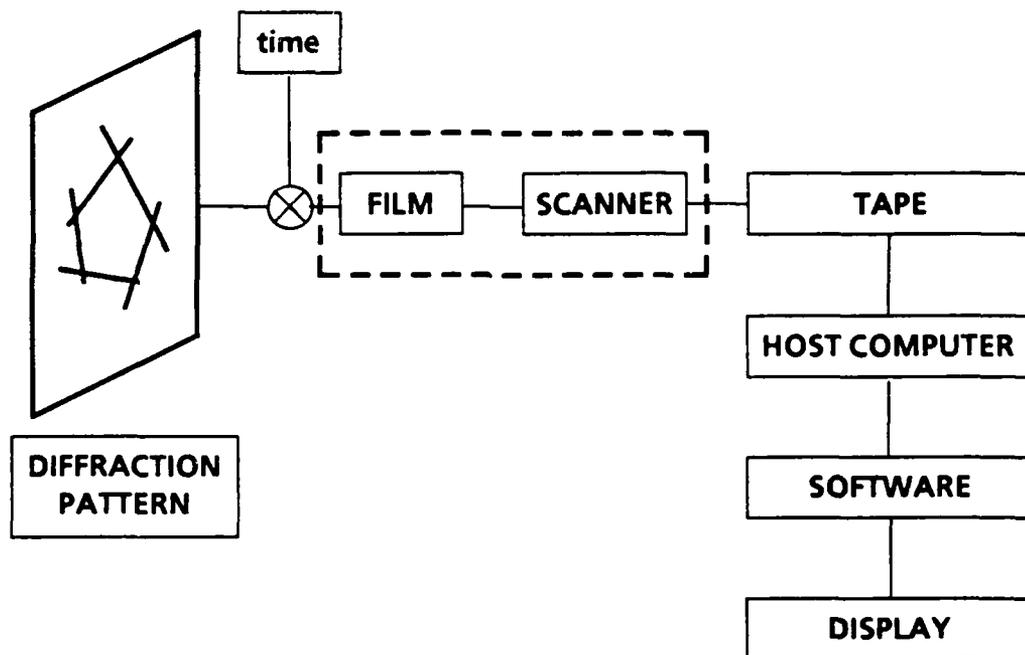
Fig. 4.20. Block diagram of experimental steps used to analyze diffraction pattern photographs. The diffraction pattern was recorded by exposing a piece of film. The optical density of the film was scanned, and the 8-bit data was analyzed using computer software.

Inc.) and specialized user programs; these manipulations enabled display and analysis of the experimental data. We describe in Appendix D some of the elements of Fig. 4.20.

Figure 4.21 shows a simplified block diagram of the experimental steps. The net effect of the film–scanner step was to apply a D-log-*e* curve to the diffraction pattern intensity. The output of the scanner was a pixel value (PV) from 0–255, however, not optical density (although there is a direct relationship). More appropriately, then, the film–scanner step applied a PV-log-*e* curve to the diffraction pattern intensity.

The data obtained by scanning and digitizing the photographs in Fig. 4.19 were sampled along a scan line corresponding to the P1–P2 scan line
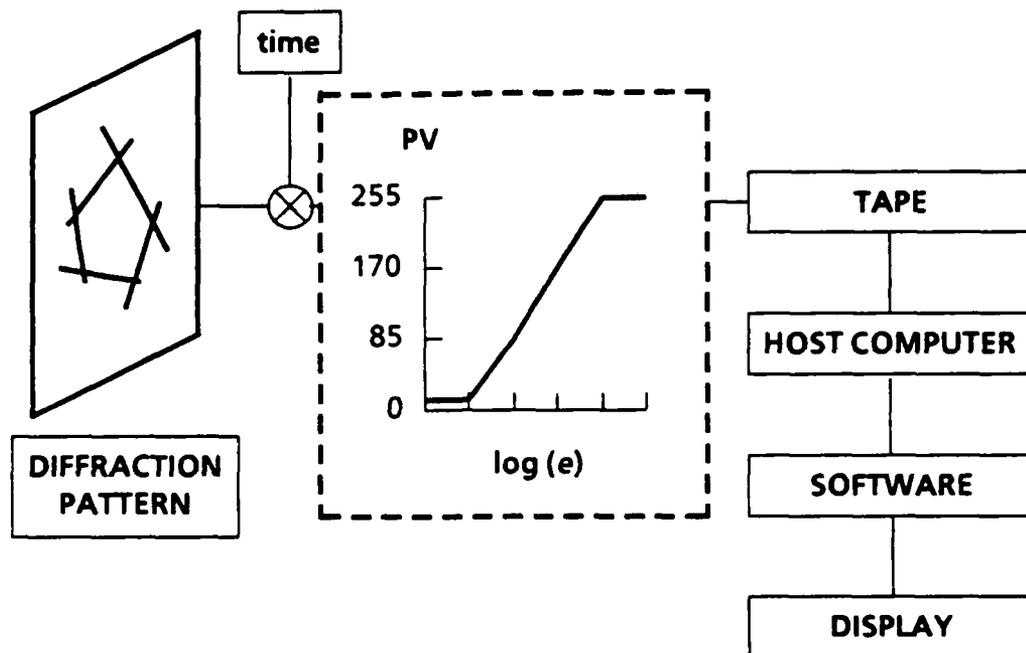
Fig. 4.21. Simplified block diagram of experimental steps used to analyze diffraction pattern photographs. Effectively, the film–scanner step applied a PV-log-e curve to the diffraction pattern intensity.

of Fig. 4.10 (–12.775 mm $\leq x \leq$ 12.775 mm, y = 0 mm). Applying Eqs. (D.2)–(D.4) and the value for the exposure time, the pixel data were converted into normalized intensity data (logarithm of intensity). The experimental results are shown in Fig. 4.22 for (a) 1/1000 second, (b) 1/60 second, (c) 1/4 second, and (d) 4 second exposures. A single frame had a response over about two and one-half orders of magnitude of intensity; the upper and lower clipping limits of (c) and (d) make this obvious. The high intensity features were recorded by the short exposures, and the low intensity features were observed in the long exposures.

We produced a composite figure by averaging the nonclipped data from each curve [see Fig. 4.23(a)]. As a result the diffraction pattern was sampled over roughly six orders of magnitude. We show in Fig. 4.23(b) the
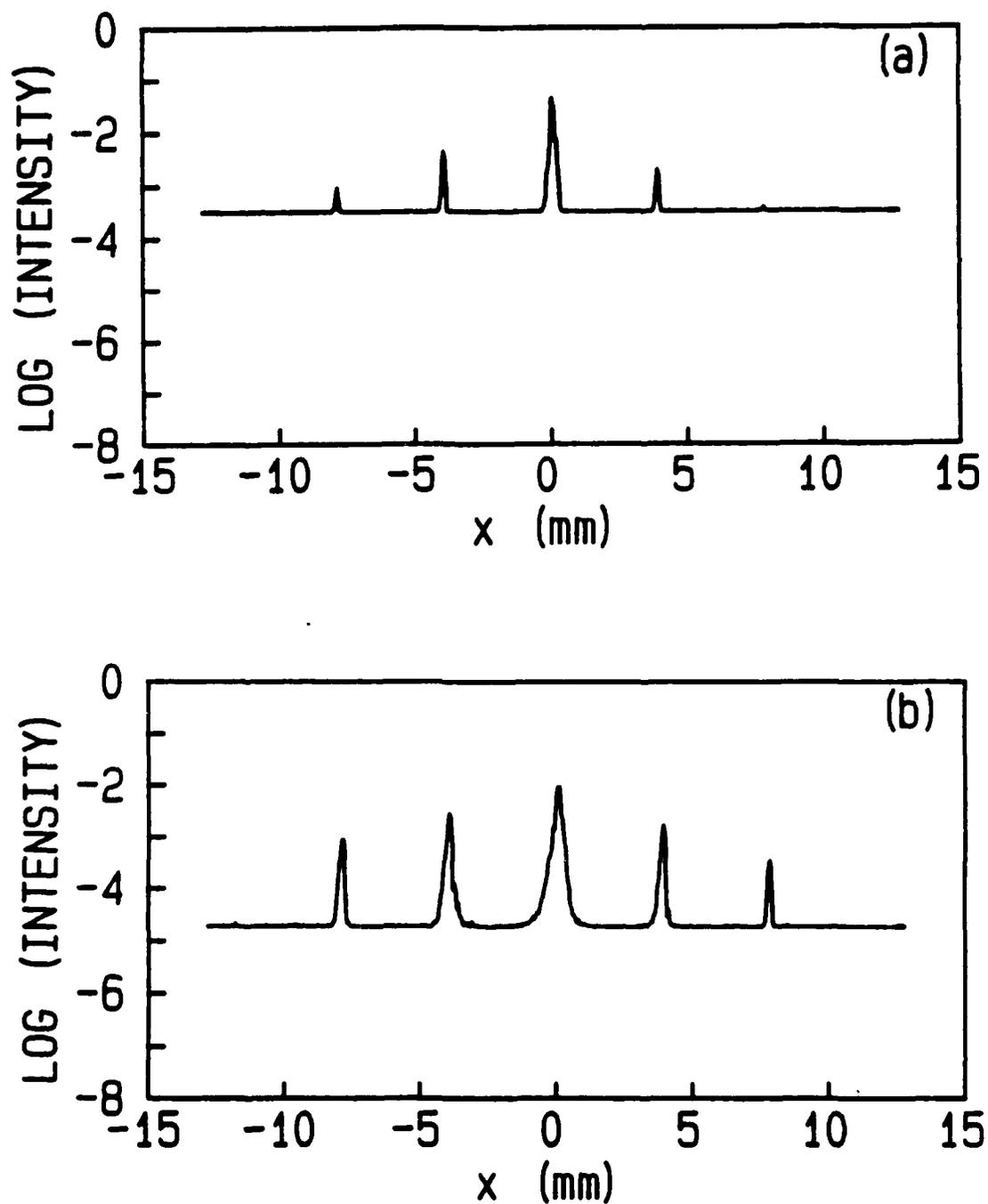
Fig. 4.22. Logarithm of normalized intensity vs position extracted from experimental photographs of optical transform of nested pentagonal aperture. The sampled data is along scan line P1-P2. Shown are various exposure times: (a) 1/1000 second and (b) 1/60 second.

148
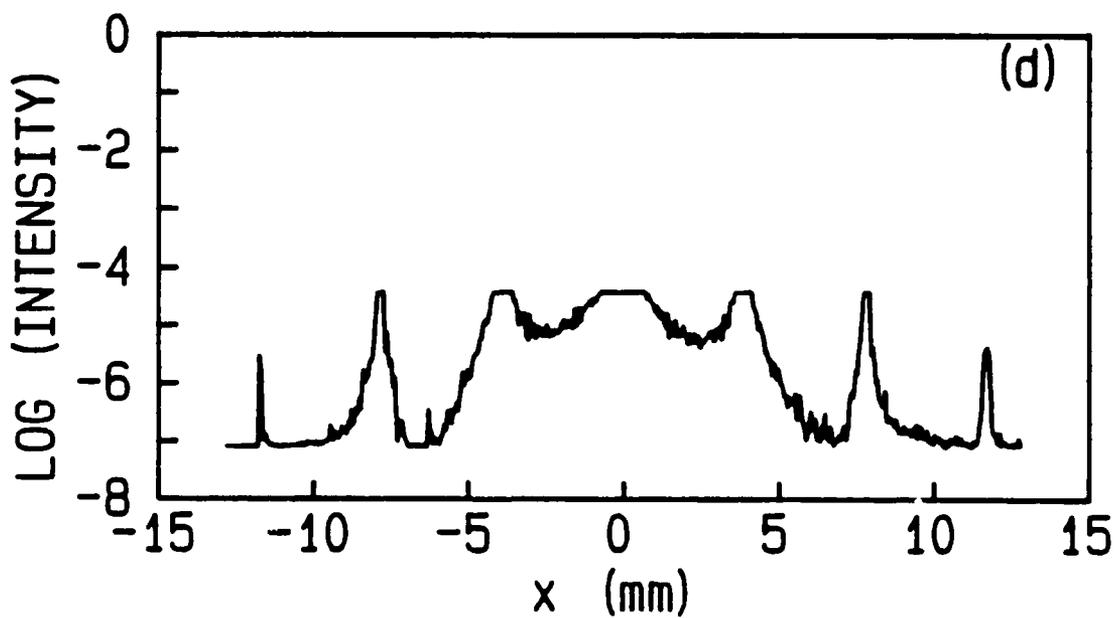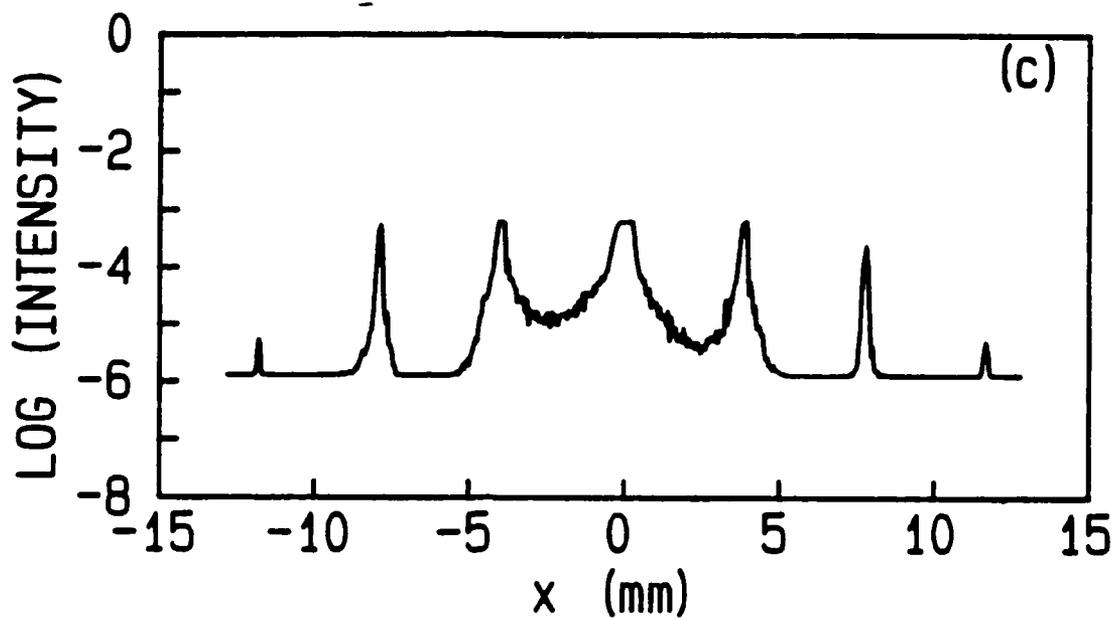
Fig. 4.22. Logarithm of normalized intensity vs position extracted from experimental photographs of optical transform of nested pentagonal aperture. The sampled data is along scan line P1-P2. Shown are various exposure times: (c) 1/4 second and (d) 4 seconds.

Fig. 4.23. Comparison of experimental data and theoretical calculations along scan line P1–P2. A composite curve derived from Fig. 4.22 is shown in (a). The theoretical prediction is plotted in (b). When the experimental data is superimposed as points (●) in (c), the agreement is seen to be very good.
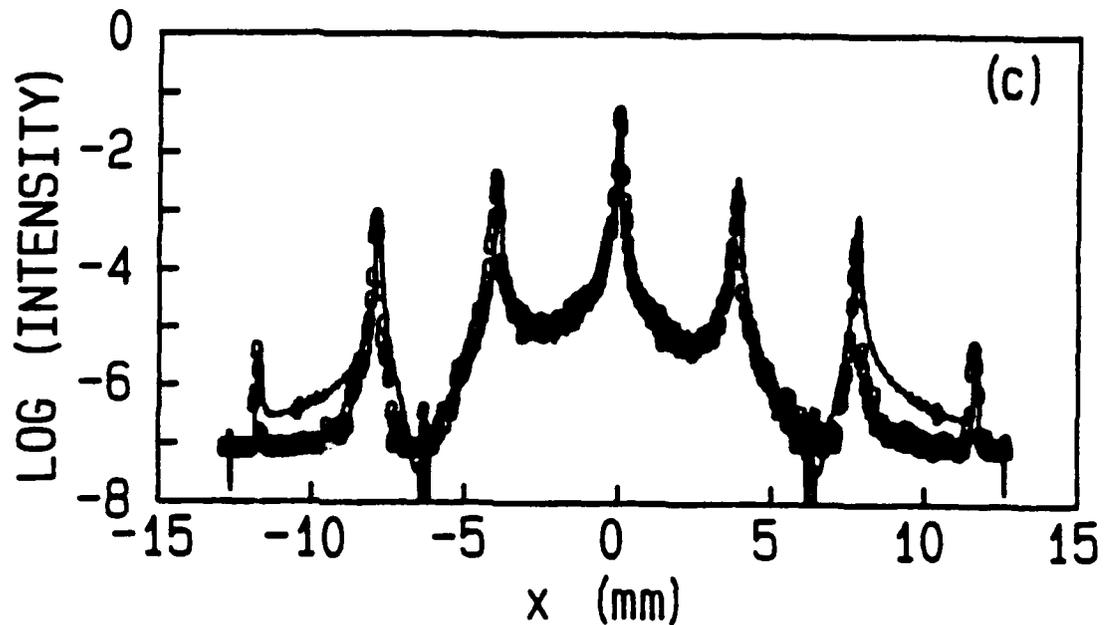
Fig. 4.23. Comparison of experimental data and theoretical calculations along scan line P1–P2. A composite curve derived from Fig. 4.22 is shown in (a). The theoretical prediction is plotted in (b). When the experimental data is superimposed as points (●) in (c), the agreement is seen to be very good.

computational result for an aperture with parameters $C = 2.1814$ mm, $d = 0.072$ mm, and $2a = 0.0475$ mm along the scan line -12.775 mm $\leq x \leq 12.775$ mm, $y = 0.05$ mm. The selected parameters are close to the measured parameters of the aperture (within the uncertainty of the measurement). Owing to the parameter sensitivity of the computational result along this scan line, we adjusted the parameters to the above values so that the correspondence in Fig. 4.23(c), which shows the experimental data as points (●), was satisfactory.

We can see that the theoretical curve closely approximates the experimental data. The height of the peaks (diffracted orders) corresponds well, and the tracking of the intensity between the zero, first, and second

orders is very good. There is, however, a deviation of the data from the theory between the second and third orders. The source of this discrepancy is unclear, but it is likely due to a mismatch of the parameters specified for the computations from their acutal values. The demonstrated sensitivity of the computational results to small changes in the parameter values makes this explanation plausible.

We extracted experimental data along paths P3–P4 and P5–P6 from the photographs in a similar manner. Composites of the data over the wide dynamic range are plotted in Fig. 4.24 (P3–P4) and Fig. 4.25 (P5–P6). The theoretical calculations fit the experimental data points (●) very well. The theory predicts the diffracted order peak intensity in both figures, and the spike line crossings in Fig. 4.25 are represented well. The slight shifts of the the computed maxima locations from the experimentally observed subsidiary maxima are due to errors in image registration and parameter specification.

Finally, we applied the PV-log-e curve, the precise form of which is given in Eq. (D.1), to the nested pentagon computational result, Eq. (14). The resulting theoretically predicted photographs can be compared visually with the experimental photographs. The procedure is equivalent to a piecewise linear mapping of the LL7 representation. We need only determine how to map pixel values in that representation (corresponding to normalized intensity of the pattern over seven orders of magnitude) to pixel values corresponding to a particular exposure time. Essentially, we want to simulate the effects of the film and exposure time.

152
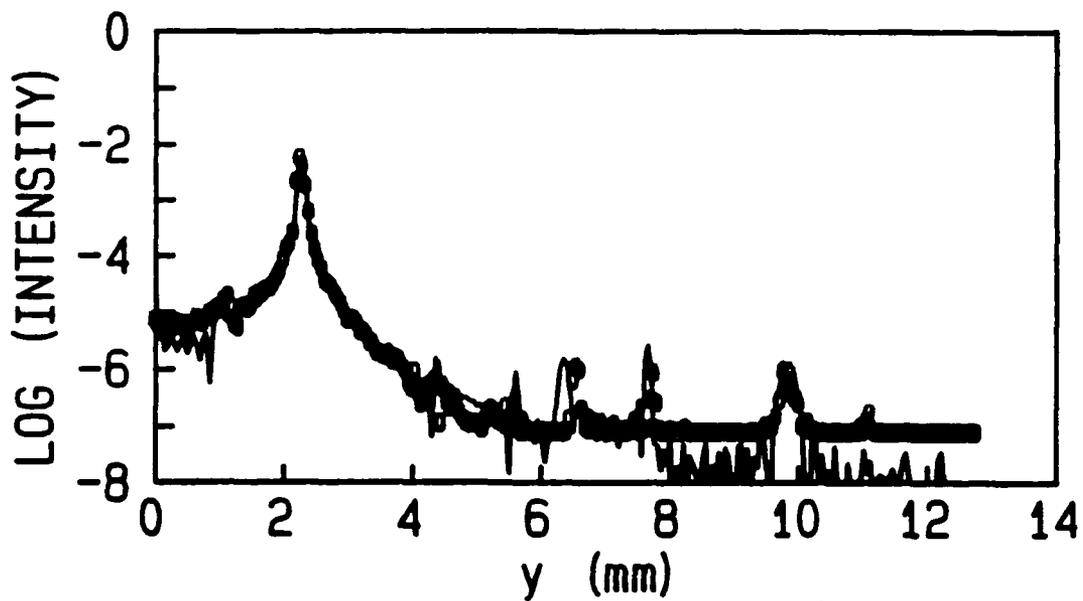
Fig. 4.24. Comparison of experimental data and theoretical calculations along scan line P3–P4. The theoretical prediction (——) agrees well with the composite curve derived from the experimental data (●).
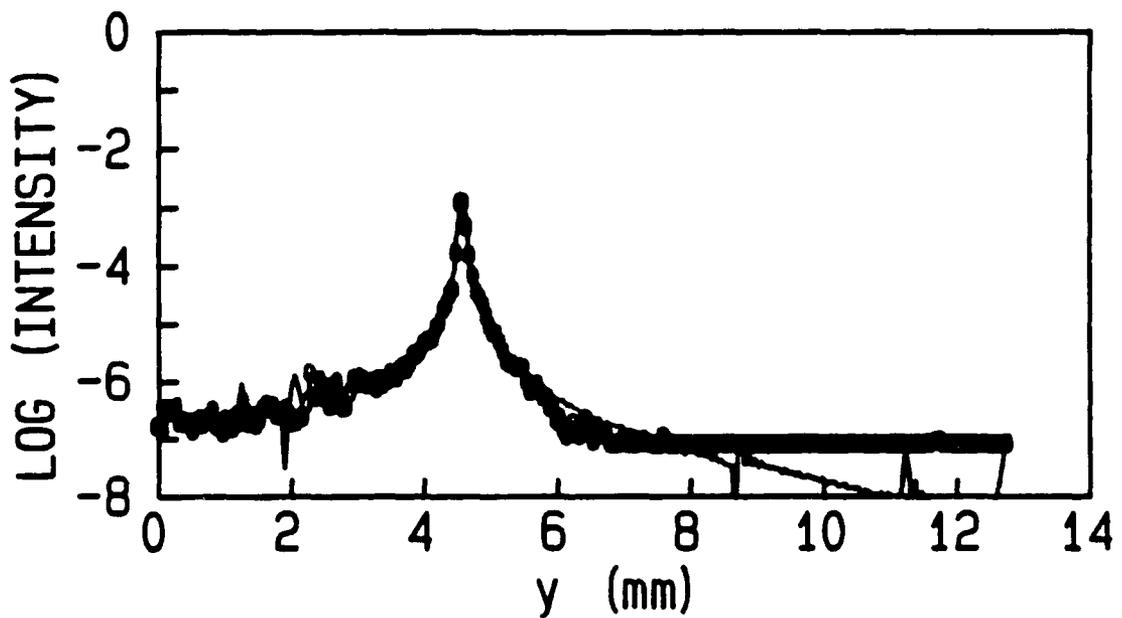
Fig. 4.25. Comparison of experimental data and theoretical calculations along scan line P5–P6. The theoretical prediction (——) agrees well with the composite curve derived from the experimental data (●).

A sample calculation will help explain the procedure. The PV-log-e curve of Fig. D.1 and Eq. (D.1) has a breakpoint in the curve at $(log\ e, PV) =$ (−3.166,13). For a 1/1000 second exposure, this $log\ e$ value was due to $log\ E_0 E_0^* = -3.486$. In the LL7 representation, this is pixel value 128 [$= 255((-3.486/7) + 1)$]. Therefore, all LL7 pixel values less than or equal to 128 must be mapped to output pixel value 13. The next breakpoint is at $(log\ e, PV) = (-1.602,126)$. The same steps yield $log\ E_0 E_0^* = -1.922$ or an LL7 pixel value of 185. Hence, LL7 pixel value 185 must be mapped to output pixel value 126.

In Table 4.1 we give the set of five breakpoint pairs necessary to transform an LL7 representation into a photograph for each exposure indicated. The first value in each pair corresponds to the original value of the LL7 input image. The second value corresponds to the output value into which the first value will be mapped. Intermediate input values are mapped linearly into output values.

We show the results in Figs. 4.26–4.29 for 1/1000 second, 1/60 second, 1/4 second, and 4 second exposure times. Shown in the figures are (a) the theoretically predicted photographs and (b) the scanned and digitized experimental photographs. The visual agreement is excellent. The high-

Table 4.1. Breakpoint Pairs for Piecewise Linear Mapping of LL7 Representation to T-MAX 100 (High Contrast) Photographs.

Exposure Time	Breakpoint Pairs
1/1000 second	(0,13) (128,13) (185,126) (226,255) (255,255)
1/60 second	(0,13) (83,13) (140,126) (181,255) (255,255)
1/4 second	(0,13) (41,13) (98,126) (139,255) (255,255)
4 seconds	(0,13) (−3,13) (54,126) (95,255) (255,255)
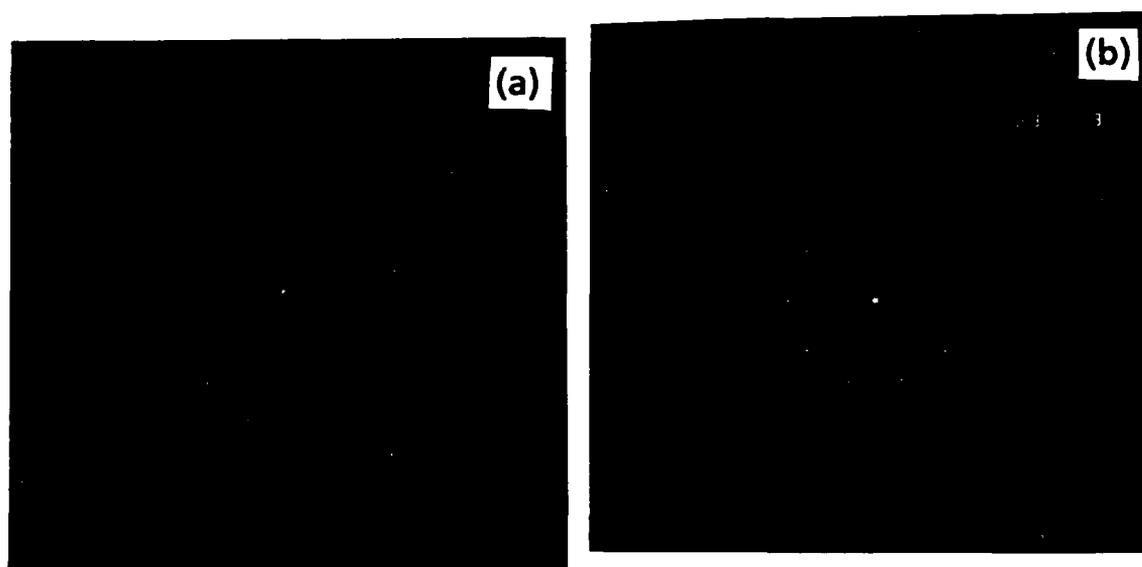
Fig. 4.26. Visual comparison of theoretically predicted photograph to scanned experimental photograph. The exposure time is 1/1000 second: (a) theory and (b) experiment.
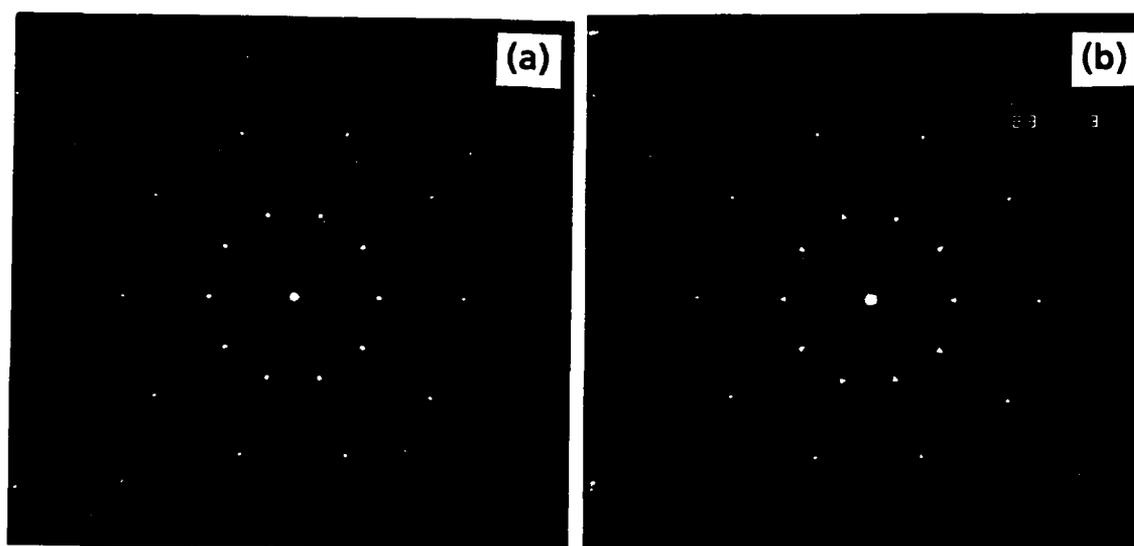
Fig. 4.27. Visual comparison of theoretically predicted photograph to scanned experimental photograph. The exposure time is 1/60 second: (a) theory and (b) experiment.
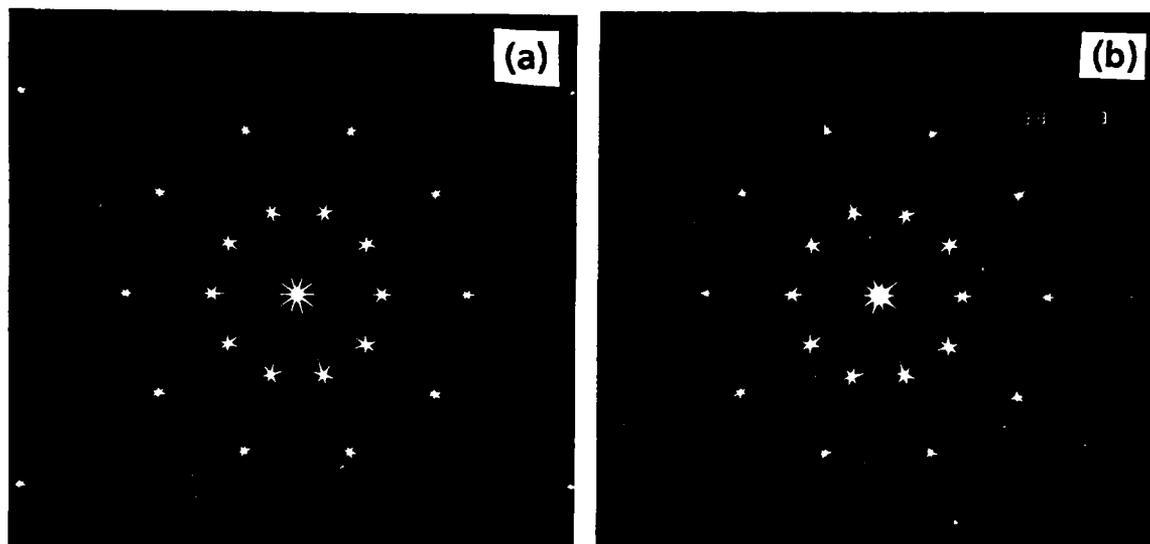
Fig. 4.28. Visual comparison of theoretically predicted photograph to scanned experimental photograph. The exposure time is 1/4 second: (a) theory and (b) experiment.
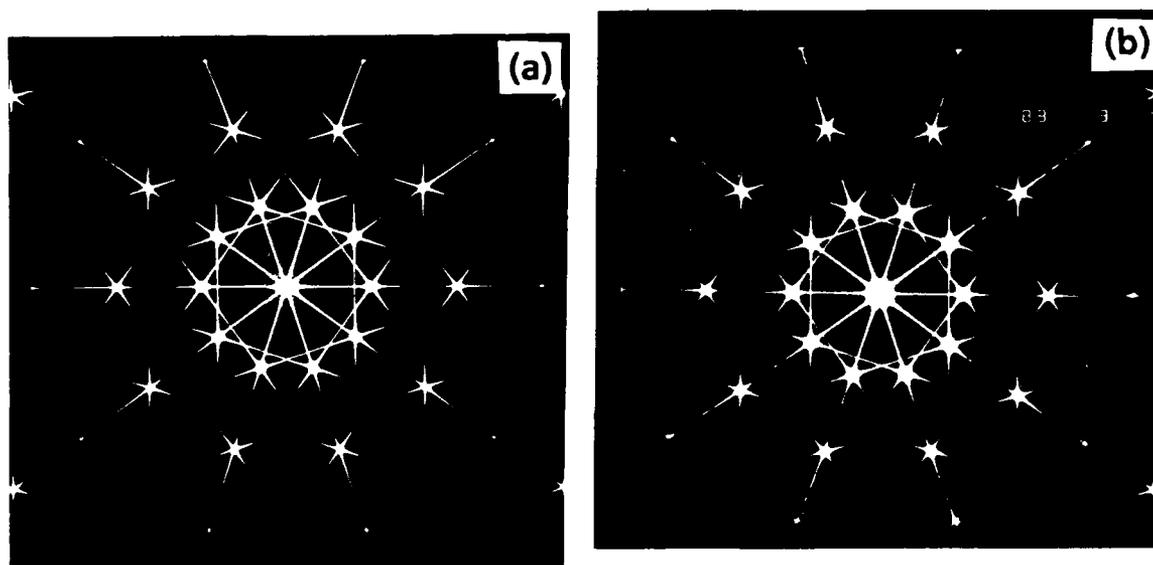
Fig. 4.29. Visual comparison of theoretically predicted photograph to scanned experimental photograph. The exposure time is 4 seconds: (a) theory and (b) experiment.

intensity features correspond well in the short exposure frames; the low-intensity details are represented accurately in the long exposure frames.

4.3 Summary

In this chapter we derived a closed-form solution for the nested polygonal aperture diffraction pattern produced in the focal plane of a converging beam. The solution predicts accurately the optical intensity measured and explains clearly the interesting low-intensity features observed in the diffraction pattern of a nested pentagonal aperture.

We began by reviewing a procedure for computing the diffraction pattern of a polygonal aperture. In that procedure one decomposes the aperture into triangular apertures [Eq. (1) and Fig. 4.2]. The diffraction patterns from these elemental triangles are added coherently to yield the solution. We illustrated the procedure by presenting computational results for a regular polygon, an octagon [Eqs. (6)–(9) and Fig. 4.4]. In this case the elemental triangles were isosceles triangles. These triangles are precisely the apertures studied in §3.2.1; we referred to them as triangular sections of a circular aperture.

In Fig. 4.5 we showed how a nested pentagonal aperture is sectioned. The elemental building blocks for nested apertures are grating-like sections. By representing a grating-like section as the product of a grating function and the section function [Eq. (10)], we were able to write a simple formula for its Fourier transform [Eq. (13)]. In Fig. 4.7 we showed how the spiked pattern of the section is replicated at the diffracted orders produced by the grating.

Then we wrote the general result for a nested polygonal aperture. The solution is a coherent sum of suitably rotated grating-like sections; we presented Eqs. (14)–(17), which constituted the result for nested pentagons. Careful analysis of the predicted optical transform pattern revealed the presence of some curious features. The most striking is that one can trace out double nested pentagons in the diffraction pattern. This feature is clearly seen in Fig. 4.8. We also showed a simple graphical representation that predicts this feature (Fig. 4.9).

In §4.2.3 we presented high-contrast photographs of the diffraction pattern recorded for a wide range of exposure times (Fig. 4.19). The photographs confirmed qualitatively the presence of the striking features predicted by our analytic solution. An additional qualitative confirmation of the analytic results was presented at the end of §4.2.4. We simulated the effect of film in recording the optical intensity and compared theoretically predicted photographs with scanned and digitized photographs.

As a quantitative test of the validity of our calculations, we measured the optical density of the film, deduced the optical intensity, and compared single line scans to theoretical calculations. The line scans were along several spikes in the diffraction pattern (see Fig. 4.10). The agreement between theory and experiment was very good over more than six orders of magnitude [Figs. 4.23(c), 4.24, and 4.25]. Thus, the representation of a nested pentagonal aperture by grating-like sections is valid and accurate. The high- and low-intensity features are accounted for by the theoretical calculations. Further, the mathematical construction of the solutions

permits us to explain the origin and behavior of these features in terms of the Fourier transform of simple elements and structures.

The treatment of different nested regular polygons, e.g., equilateral triangle, square, and hexagon, and so on, is similar. One simply uses the correct opening angle in the section transform and sums the appropriate number of sections, rotated to the corresponding positions. And the extension to nested arbitrary polygons is straightforward. The nested polygon is decomposed into elemental grating-like triangles. The transform of each elemental triangle is slightly more complicated than the section transform, but still quite manageable.

REFERENCES

M. Abramowitz and I. Stegun, Eds., *NBS Handbook of Mathematical Functions*, (U. S. Government Printing Office, Washington, DC, 1972).

W. Andrejewski, "Die Beugung elektromagnetischer Wellen an der leitenden Kreisscheibe und an der kreisförmigen Öffnung im leitenden ebenen Schirm," Ph. D. dissertation, University of Aachen, 1953.

C. L. Andrews, "Near-field diffraction, experimental and theoretical," J. Opt. Soc. Am. **72**, 1825 (1982).

C. L. Andrews and I. Chung, "Diffraction on the axes of disks and apertures," J. Opt. Soc. Am. **70**, 813–816 (1980).

E. Arvas and R. F. Harrington, "Computation of the magnetic polarizability of conducting disks and the electrical polarizability of apertures," IEEE Trans. Antennas Propagat. AP-31, 719–724 (1983).

B. B. Baker and E. T. Copson, *The Mathematical Theory of Huygens' Principle* (Clarendon Press, Oxford, 1953).

M. J. Bastiaans, "Gabor's signal expansion and degrees of freedom of a signal," Opt. Act. **29**, 1223–1229 (1982).

H. A. Bethe, "Theory of diffraction by small holes," Phys. Rev. **66**, 163–182 (1944).

A. J. Bogush, Jr. and R. E. Elkins, "Gaussian field expansions for large aperture antennas," IEEE Trans. Antennas Propagat., AP-34, 228–243 (1986).

M. Born and E. Wolf, *Principles of Optics*, 6th ed. (Pergamon, New York, 1980).

C. J. Bouwkamp, "A note on singularities occurring at sharp edges in electromagnetic theory," Physica 12, 467–474 (1946); Math. Rev. 8, 179–180 (1947); "On Bethe's theory of diffraction by small holes," Philips Res. Rep. 5, 321–332 (1950a); "On the diffraction of electromagnetic waves by small circular disks and holes," Philips Res. Rep. 5, 401–422 (1950b); "Diffraction theory: A critique of some recent developments," New York University Mathematics Research Group, Research Report No. EM-50 (1953); "Diffraction Theory," Rept. Progr. Phys. 17, 35–100 (1954); "Theoretical and numerical treatment of diffraction through a circular aperture," IEEE Trans. Antennas Propagat. AP-18, 152–176 (1970).

C. J. Bouwkamp and N. G. de Bruijn, "The problem of optimum antenna current distribution," Philips Res. Rep. 1, 135–138 (1946).

G. D. Boyd and J. P. Gordon, "Confocal multimode resonator for millimeter through optical wavelength masers," Bell Sys. Tech. J. 40, 489–508 (1961).

G. D. Boyd and H. Kogelnik, "Generalized confocal resonator theory," Bell Sys. Tech. J. 41, 1347–1369 (1962).

R. N. Bracewell, *The Fourier Transform and Its Applications* (McGraw-Hill, New, 1978).

W. Braunbek, "On the diffraction field near a plane-screen corner," IRE Trans. Antennas Propagat. AP-4, 219–223 (1956).

C. M. Butler, "A review of electromagnetic diffraction by small apertures in conducting surfaces," Proc. International Symposium Electromagnetic Compatibility, IEEE 13, 86–91 (1978).

C. M. Butler, Y. Rahmat-Samii, and R. Mittra, "Electromagnetic penetration through apertures in conducting surfaces," IEEE Trans. Electromagnetic Compat. EMC-20, 82–93 (1978).

M. Cagnet, M. Francon, and J. C. Thrierr, *Atlas of Optical Phenomena* (Springer-Verlag, Berlin and Prentice-Hall, Englewood Cliffs, New Jersey, 1962).

C. R. Carpenter, "Electromagnetic diffraction by a thin conducting disk on the shadow side in the near-region," Ph. D. dissertation, State University of New York at Albany, 1973.

D. Casasent and T. Luu, Appl. Opt. 17, 1701–1708 (1978); Appl. Opt. 17, 2973–2980 (1978).

I. Chung, C. L. Andrews, and L. F. Libelo, "Near-field diffraction on axes of disks," J. Opt. Soc. Am. 67, 1561–1566 (1977).

S. B. Cohn, "Determination of aperture parameters by electrolytic-tank measurements," Proc. IRE 39, 1416–1421 (1951); "The electric polarizability of apertures of arbitrary shape," Proc. IRE 40, 1069–1071 (1952).

E. T. Copson, "An integral-equation method of solving diffraction problems," Proc. Roy. Soc. (London) A186, 100–118 (1946); "Diffraction by a plane screen," Proc. Roy. Soc. (London) A202, 277–284 (1950); Math. Rev. 12, 774 (1951).

J. C. Dainty, *Physical Optics I*, lecture notes for Physical Optics 461, Fall 1982 (The Institute of Optics, Rochester, New York, 1983).

F. De Meulenaere and J. Van Bladel, "Polarizability of some small apertures," IEEE Trans. Antennas Propagat. AP-25, 198–205 (1977).

R. De Smedt and J. Van Bladel, "Magnetic polarizability of some small apertures," IEEE Trans. Antennas Propagat. AP-28, 703–707 (1980); "Field singularities at the tip of a metallic cone of arbitrary cross section," IEEE Trans. Antennas Propagat. AP-34, 865–870 (1986); "Field singularities near aperture corners," Inst. Elec. Eng. Proc. 134 A, 694–698 (1987).

W. H. Eggimann, "Higher-order evaluation of electromagnetic diffraction by circular disks," IRE Trans. Microwave Theory Tech. MTT-9, 408–418 (1961).

P. D. Einziger, S. Raz, and M. Shapira, "Gabor representation and aperture theory," J Opt. Soc. Am. A 3, 508–522 (1986).

R. E. English, Jr. and N. George, "Diffraction under Gaussian illumination: a Gaussian beam expansion approach," 1986 Annual Meeting of The Optical Society of America in Seattle, WA.

R. E. English, Jr. and N. George, "Diffraction from a small square aperture: approximate aperture fields," J. Opt. Soc. Am. A 5, 192–199 (1988).

F. D. Feiock, "Wave propagation in optical systems with large apertures," J. Opt. Soc. Am. 68, 485–489 (1978).

C. Flammer, "The vector wave function solution of the diffraction of electromagnetic waves by circular disks and apertures. I. Oblate spheroidal vector wave functions," J. Appl. Phys. 24, 1218–1223 (1953a); "The vector wave function solution of the diffraction of electromagnetic waves by circular disks and apertures. II. The diffraction problems," J. Appl. Phys. 24, 1224–1231 (1953b).

A. G. Fox and T. Li, "Resonant modes in a maser interferometer," Bell Sys. Tech. J. 40, 453–488 (1961).

D. Gabor, "Theory of communication," J. Inst. Elec. Eng. (London) 93III, 429–457 (1946).

S. Ganci, "Simple derivation of formulas for Fraunhofer diffraction at polygonal apertures from Maggi-Rubinowicz transformation," J. Opt. Soc. Am. A 1, 559–561 (1984).

R. R. Goldberg, *Fourier Transforms* (Cambridge Univesity Press, Cambridge, 1965).

J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1968).

G. Goubau and F. Schwering, "On the guided propagation of electromagnetic beam waves," IRE Trans. Antennas Propagat. **AP-9**, 248–256 (1961).

I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, ed. by A. Jeffrey (Academic, New York, 1980).

G. Harburn, C. A. Taylor, and T. R. Welberry, *Atlas of Optical Transforms* (Cornell University Press, Ithaca, 1975).

R. F. Harrington, *Field Computation by Moment Methods* (Macmillan, New York, 1968).

F. S. Harris, Jr., "Light diffraction patterns," Appl. Opt. **3**, 909–913 (1964).

J. C. Heurtley, "Hyperspheroidal functions – optical resonators with circular mirrors," in *Proceedings of the Symposium on Quai-Optics* (Polytech. Press, Brooklyn, 1964) pp. 367–375.

J. C. Heurtley and W. Streifer, "Optical resonator modes – circular reflectors of spherical curvature," J. Opt. Soc. Am. **55**, 1472–1479 (1965).

IMSL Library, Ed. 9, Copyright 1982 by IMSL, Inc.

International Imaging Systems, Inc., 1500 Buckeye Drive, Milpitas, California.

J. D. Jackson, *Classical Electrodynamics*, 2nd ed. (John Wiley, New York, 1975).

D. L. Jaggard and C. H. Papas, "On the application of symmetrization to the transmission of electromagnetic waves through small convex apertures of arbitrary shape," Appl. Phys. **15**, 21–25 (1978).

R. W. James, *The Optical Principles of the Diffraction of X-Rays* (G. Bell and Sons, London, 1948).

D. S. Jones, "Note on diffraction by an edge," Quart. Jour. Mech. Appl. Math. **3**, 420–434 (1950); "Diffraction by an edge and by a corner," Quart. Jour. Mech. Appl. Math. **5**, 363–378 (1952).

D. Joyeux and S. Lowenthal, Appl. Opt. **21**, 4368–4372 (1982).

P. Kane, "Diffraction by elliptical apertures: Theory and experiment," M. S. Thesis, University of Rochester, Rochester, New York, 1984.

R. B. Kieburtz, "Electromagnetic Diffraction by a Periodically Apertured Conducting Screen," Ph.D. dissertation, University of Washington, Seattle, Washington, 1961.

R. B. Kieburtz and A. Ishimaru, "Aperture fields of an array of rectangular apertures," IRE Trans. Antennas Propagat. **AP-10**, 663–671 (1962).

R. F. Kleinman, "Low frequency electromagnetic scattering," in *Electromagnetic Scattering*, ed. by P. L. E. Uslenghi (Academic, New York, 1978) pp. 1–28.

Eastman Kodak Co., Rochester, New York; *Kodak Plates and Films for Science and Industry*, Kodak Data Book P-9 (1962); *Characteristics of Kodak Plates for Scientific and Technical Applications*, Kodak Data Book P-140 (19xx); *Kodak TMAX Professional Films*, pamphlet F-25 (1987); Kodak, T-Max, HC-110, and Photo-Flo are trademarks of Eastman Kodak Co.

H. Kogelnik, "On the propagation of Gaussian beams of light through lenslike media including those with a loss or gain variation," Appl. Opt. **4**, 1562–1569 (1965).

H. Kogelnik and T. Li, "Laser beams and resonators," Appl. Opt. **5**, 1550–1567 (1965).

J. Komrska, "Fraunhofer diffraction at apertures in the form of regular polygons. I.," Opt. Act. **19**, 807–816 (1972); "Fraunhofer diffraction at apertures in the form of regular polygons. II.," Opt. Act. **20**, 549–563 (1973); "Simple derivation of formulas for Fraunhofer diffraction at polygonal apertures," J. Opt. Soc. Am. **72**, 1382–1384 (1982).

J. A. Kong, *Electromagnetic Wave Theory* (John Wiley, New York, 1986) §5.8, pp. 436–443.

S. H. Lee, Ed., *Optical Information Processing* (Springer-Verlag, New York, 1981) ch. 1.

R. L. Liboff, *Introductory Quantum Mechanics* (Holden-Day, San Francisco, 1980) §7.3, pp. 187–195.

A. C. Livanos and N. George, "Edge diffraction of a convergent wave," Appl. Opt. **14**, 608–613 (1975).

J. J. Maciel and L. B. Felsen, "On Gaussian discretization of radiation from distributed aperture fields," 1987 IEEE Antennas and Propagation Society International Symposium, Symposium Digest, vol. 1, 404–407 (June 15-19, 1987).

W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer-Verlag, New York, 1966) §5.5, 5.6, pp. 239–255.

MASSCOMP, One Technology Park, Westford, Massachusetts.

J. R. Mautz and R. F. Harrington, "Electromagnetic transmission through a rectangular aperture in a perfectly conducting screen," Dept. Elect. Comput. Eng., Syracuse University, Tech. Rep. TR-76-1, February, 1976.

C. E. K. Mees, *The Theory of the Photographic Process*, rev. ed. (Macmillan, New York, 1954).

V. J. Meixner, "Die Kantenbedingung in der Theorie der Beugung Elektromagnetischer Wellen an Volkommen Leitenden Ebenen Schirm," Ann. Phys. **6**, 2–9 (1949); "The behavior of electromagnetic fields at edges," IEEE Trans. Antennas Propagat. AP-21, 442–446 (1972).

V. J. Meixner and W. Andrejewski, "Strenge Theorie der Beugung ebener elektromagnetischer Wellen an der vollkommen leitenden Kreisscheibe und an der Kreisförmigen Öffnung in vollkommen leitenden ebenen Schirm," Ann. Physik **7**, 157–168 (1950).

R. F. Millar, "A note on diffraction by an infinite slit," Can. J. Phys. **38**, 38–47 (1960).

R. Mittra et al., "A new look at the thin-plate scattering problem," Radio Sci. **8**, 869–875 (1973).

Y. Nomura and S. Katsura, "Diffraction of electromagnetic waves by circular plate and circular hole," J. Phys. Soc. Japan **10**, 285–304 (1955).

E. E. Okon and R. F. Harrington, "The polarizabilities of electrically small apertures of arbitrary shape," IEEE Trans. Electromagnetic Compat. EMC-23, 359–366 (1981).

Optronics International, Inc., Chelmsford, Massachusetts; *Operation and Maintenance Manual*, Rev. D, (19 March 1984).

C. H. Papas, *Theory of Electromagnetic Wave Propagation* (McGraw-Hill, New York, 1965) §3.5, pp. 50–56.

C. L. Pekeris, "Comments on Bethe's theory of diffraction of electromagnetic waves by small holes," Phys. Rev. **66**, 351 (1944).

W. Pogorzelski, *Integrals Equations and the Applications*, vol. 1 (Pergamon, Oxford, 1966) pp. 342–345.

Y. Rahmat-Samii and R. Mittra, "Integral equation solution and RCS computation of thin rectangular plate," IEEE Trans. Antennas Propagat. AP-22, 608–610 (1974); "Electromagnetic coupling through small apertures in a conducting screen," IEEE Trans. Antennas Propagat. AP-25, 180–187 (1977).

S. M. Rao, D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," IEEE Trans. Antennas Propagat. AP-30, 409–418 (1982).

Lord Rayleigh, "On the passage of waves through apertures in plane screens, and allied problems," Phil. Mag. 43, 259–272 (1897a); "On the incidence of aerial and electric waves upon small obstacles in the form of ellipsoids or elliptic cylinders, and on the passage of electric waves through a circular aperture in a conducting screen," Phil. Mag. 44, 28–52 (1897b).

T. K. Sarkar, E. Arvas, and S. M. Rao, "Application of the fast Fourier transform and the conjugate gradient method for efficient solution of electromagnetic scattering from both electrically large and small conducting bodies," Electromagnetics 5, 99–122 (1985).

R. M. Sillitto and W. Sillitto, "A simple Fourier approach to Fraunhofer diffraction by triangular apertures," Opt. Act. 22, 999–1010 (1975).

W. Sillitto, "Fraunhofer diffraction at straight-edged apertures," J. Opt. Soc. Am. 69, 765–770 (1979).

R. C. Smith and J. S. Marsh, "Diffraction patterns of simple aperture," J. Opt. Soc. Am. 64, 798–803 (1974).

F. Smithies, *Integral Equations* (Cambridge University Press, London, 1958).

W. R. Smythe, "The double current sheet in diffraction," Phys. Rev. 72, 1066–1070 (1947); *Static and Dynamic Electricity*, 3rd ed. (McGraw-Hill, New York, 1968).

A. Sommerfeld, Math. Ann. 47, 317 (1896); *Optics* , trans. by O. LaPorte and P. A. Moldauer (Academic, New York, 1954).

A. F. Stevenson, "Solution of electromagnetic scattering problems as power series in the ratio (dimension of scatterer)/wavelength," J. Appl. Phys. 24, 1134–1142 (1953).

J. M. Stone, *Radiation and Optics* (McGraw-Hill, New York, 1963).

W. Streifer, "Optical resonator modes – rectangular reflectors of spherical curvature," J. Opt. Soc. Am. 55, 868–877 (1965).

W. Streifer and H. Gamo, "On the Schmidt expansion for optical resonator modes," in *Proceedings of the Symposium on Quasi-Optics* (Polytech. Press, Brooklyn, 1964) pp. 351–365.

M. Suzuki, "Diffraction of plane electromagnetic waves by a rectangular aperture," IRE Trans. Antennas Propagat. AP-4, 149–157 (1956).

K. Tanaka, N. Saga, and H. Mizokami, "Field spread of a diffracted Gaussian beam through a circular aperture," Appl. Opt. 24, 1102–1106 (1985).

K. Tanaka, K. Yoshida, and M. Taguchi, "Analytical and experimental investigations of the diffraction field of a Gaussian beam through a sequence of apertures: applicability of the beam mode expansion method," Appl. Opt. 27, 1310–1312 (1988).

Von H. J. Tiziani and W. Witz, "Anwedung des Schwärzungsreliefs in opto-elektronischem Formerkennungs-System," Optik 32, 484–495 (1971).

L. A. Vainshtein, "Open resonators for lasers," Zh. Eksperim. i Teor. Fiz. (USSR) 44, 1050–1067 (1963) trans. in Soviet Phys. – JETP 17, 709–719 (1963).

J. Van Bladel, *Electromagnetic Fields* (McGraw-Hill, New York, 1964); "The multipole expansion revisited," Arch. Elek. Übertrag. 31, 407–411 (1977); "Field penetration through small apertures: The first order correction," Radio Sci. 14, 319–331 (1979).

E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, 4th ed. (Cambridge University Press, Cambridge, 1927).

E. E. Witmer, "Fraunhofer diffraction phenomena for a general polygonal aperture," Phil. Mag. 1, 132–138 (1926).

W. Witz, "Fraunhofer diffraction pattern of an aperture," J. Opt. Soc. Am. 65, 1077–1078 (1975).

APPENDIX A

TRANSFORMATION OF ELLIPTIC INTEGRALS INTO
STANDARD FORM

A general procedure for transforming an elliptic integral into
standard form is described in the book by Whittaker and Watson (1927,
Chapter 22). An elliptic integral is

$$\int r(x,y)\, dx \quad , \tag{1}$$

where r is any rational function of x and y, and y^2 is a cubic or quartic
polynomial of x. (It is assumed that r contains an odd power of y and that y^2
has no repeated factors.) We used elliptic integrals to evaluate Eq. (2.27);
the results are Eqs. (2.29) and (2.30). We demonstrate here the
transformation procedure by the step-by-step derivation of Eq. (2.30a).

Evaluate the following integral:

$$I_2 = \int_{-1}^{1} \frac{1}{(1-\xi')^{1/2}} \frac{d\xi'}{\left[(\xi-\xi')^2 + (\eta-\eta')^2\right]^{1/2}} \quad , \tag{2}$$

where $|\xi| < 1$ and $|\eta| < 1$.

First introduce some simplifying notation.

$$I_2 = \int_{-1}^{1} \frac{d\xi'}{w} \quad , \tag{3}$$

where

$$w^2 = S_1 S_2 \quad , \tag{4}$$

$$S_1 = \xi'^2 - 2\xi\xi' + \xi^2 + \upsilon^2 \quad , \tag{5}$$

$$S_2 = -\xi'^2 + 1 \quad . \tag{6}$$

167

$$o^2 = (\eta - \eta')^2 \quad . \tag{7}$$

The initial goal is to express w as the product of sums of squares, i.e.,

$$S_1 = A_1 (\xi' - a)^2 + B_1 (\xi' - \beta)^2 \quad , \tag{8}$$

$$S_2 = A_2 (\xi' - a)^2 + B_2 (\xi' - \beta)^2 \quad . \tag{9}$$

We achieve this by finding the root λ such that $S_1 - \lambda S_2$ is a perfect square.

$$S_1 - \lambda S_2 = \xi'^2 (1 + \lambda) + \xi'(-2\xi) + (\xi^2 + o^2 - \lambda) \quad . \tag{10}$$

There are two roots:

$$\lambda_{1,2} = \frac{1}{2} \left\{ [\xi^2 + o^2 - 1] \pm [(\xi^2 + o^2 - 1)^2 + 4o^2]^{1/2} \right\} \quad . \tag{11}$$

The perfect squares are

$$S_1 - \lambda_1 S_2 = (1 + \lambda_1)(\xi' - a)^2 \quad , \tag{12}$$

where

$$a = \frac{\xi}{1 + \lambda_1} \quad , \tag{13}$$

and

$$S_1 - \lambda_2 S_2 = (1 + \lambda_2)(\xi' - \beta)^2 \quad , \tag{14}$$

where

$$\beta = \frac{\xi}{1 + \lambda_2} \quad . \tag{15}$$

Solving Eqs. (12) and (14) for S_1 and S_2 gives Eqs. (8) and (9) where

$$A_1 = \frac{\lambda_2 (1 + \lambda_1)}{\lambda_2 - \lambda_1} \quad , \tag{16}$$

$$B_1 = -\frac{\lambda_1 (1 + \lambda_2)}{\lambda_2 - \lambda_1} \quad , \tag{17}$$

$$A_2 = \frac{1 + \lambda_1}{\lambda_2 - \lambda_1} \quad , \tag{18}$$

$$B_2 = -\frac{1+\lambda_2}{\lambda_2-\lambda_1} \quad . \tag{19}$$

Using Eqs. (3), (8), and (9), we write w^2 in terms of a new variable.

An intermediate step is

$$w^2 = S_1 S_2 = (\xi'-\beta)^4 [(A_1 t^2 + B_1)(A_2 t^2 + B_2)] \quad , \tag{20}$$

where

$$t = \frac{\xi'-\alpha}{\xi'-\beta} \quad . \tag{21}$$

One can easily show that

$$\xi'-\beta = -\left(\frac{\alpha-\beta}{t-1}\right) \quad . \tag{22}$$

Thus, Eq. (3) becomes

$$I_2 = \int_{t(\xi'=-1)}^{t(\xi'=1)} \frac{1}{\alpha-\beta} \frac{dt}{[(A_1 t^2 + B_1)(A_2 t^2 + B_2)]^{1/2}} \tag{23}$$

with the limits to be determined.

It is helpful to see that

$$(1+\lambda_1)(1+\lambda_2) = \xi^2 \tag{24}$$

and

$$\alpha\beta = 1 \quad . \tag{25}$$

Therefore, using Eq. (21),

$$t(\xi'=\pm1) = \frac{\pm1-\alpha}{\pm1-\beta} = \left(\frac{\pm1-\alpha}{\pm1-\beta}\right)\left(\frac{\alpha}{\alpha}\right) = \alpha\left(\frac{\pm1-\alpha}{\pm\alpha-\alpha\beta}\right) = \mp\alpha \quad ; \tag{26}$$

thus the limits in Eq. (23).are determined.

The final steps require another variable substitution. Some simplifying notation helps.

$$q = [(A_1 t^2 + B_1)(A_2 t^2 + B_2)]^{1/2} = (B_1 B_2)^{1/2} [(1+\frac{A_1}{B_1} t^2)(1+\frac{A_2}{B_2} t^2)]^{1/2} \quad . \tag{27}$$

Noting that $\lambda_1 > 0$ and $\lambda_2 < 0$, one can see that $A_1 > 0$ and $A_2 < 0$. Further,

$1 + \lambda_2 > 0$ for $\xi \neq 0$ and $1 + \lambda_2 = 0$ for $\xi = 0$. Therefore, $B_1 > 0$ and $B_2 > 0$.

Define

$$\gamma_1^2 = \frac{A_1}{B_1} > 0 \qquad (28)$$

and

$$\gamma_2^2 = -\frac{A_2}{B_2} > 0 \quad . \qquad (29)$$

Thus, Eq. (27) becomes

$$q = (B_1 B_2)^{1/2} [(1 + \gamma_1^2 t^2)(1 - \gamma_2^2 t^2)]^{1/2} \quad . \qquad (30)$$

The needed variable substitution is

$$\cos\phi = \gamma_2 t \quad . \qquad (31)$$

Some algebraic manipulation yields

$$I_2 = \int_{\phi(t=a)}^{\phi(t=-a)} -\frac{1/\gamma_2}{(B_1 B_2)^{1/2}(a-\beta)\left(1 + \frac{\gamma_1^2}{\gamma_2^2}\right)^{1/2}} \frac{d\phi}{[1 - k^2 \sin^2\phi]^{1/2}} \quad , \qquad (32)$$

where

$$k^2 = \frac{\gamma_1^2/\gamma_2^2}{1 + \gamma_1^2/\gamma_2^2} = \frac{\gamma_1^2}{\gamma_1^2 + \gamma_2^2} = \frac{\lambda_2}{\lambda_2 - \lambda_1} \quad . \qquad (33)$$

Equation (32) is in the standard form for an elliptic integral of the first kind.

Tedious algebra completes the derivation and shows that I_2 is expressible in terms of a complete elliptic integral of the first kind.

$$I_2 = \int_0^\pi \frac{1}{(\lambda_1 - \lambda_2)^{1/2}} \frac{d\phi}{[1 - k^2 \sin^2\phi]^{1/2}} = \frac{2}{(\lambda_1 - \lambda_2)^{1/2}} \int_0^{\pi/2} \frac{d\phi}{[1 - k^2 \sin^2\phi]^{1/2}} \quad , \qquad (34)$$

$$I_2 = \frac{2}{(\lambda_1 - \lambda_2)^{1/2}} K(k) \quad , \qquad (35)$$

where

$$\mathbf{K}(k) = \int_0^{\pi/2} \frac{d\phi}{(1 - k^2 \sin^2 \phi)^{1/2}} \qquad . \tag{36}$$

Equation (35) is Eq. (2.30a).

APPENDIX B

NUMERICAL INTEGRATION AND DIFFERENTIATION

We performed the numerical calculations of Chapter 2 in double
precision on a DEC VAX 11/750 minicomputer. The integration of Eqs.
(2.31a)–(2.31c) used the cautious adaptive Romberg extrapolation routine
(DCADRE) from the IMSL Library (1982). The numerical results
presented in Fig. 2.3 are accurate to seven or eight decimal places. The
initial calculation of Eqs. (2.31a)–(2.31c) at some sample points were only
accurate to five or six decimal places. This degradation of accuracy
introduced noticeable numerical noise in the subsequent derivative
calculations. This degradation resulted from the manner in which the
integration routine handled the integrable singularity at the point $\eta' = \eta$
in Eqs. (2.31a)–(2.31c). Note that $\sigma^2 = 0$ at this point. This type of
singularity is commonly encountered when evaluating diffraction
integrals at points within the aperture, i.e., $(x,y) \in A, z = 0$.

A small adjustment to the value of σ^2 greatly stabilized the
integration algorithm. For example, if we shifted the sample point by $\Delta\eta$
$= 10^{-7}$ (η is shifted to $\eta + \Delta\eta$), we sometimes obtained an acceptably
accurate value. A second adjustment technique involved rewriting σ^2 as

$$\sigma^2 = (\eta - \eta')^2 + z^2 \quad .$$

This permitted us to evaluate $F_{1y}{}^A$ for $z \neq 0$. A shift from $z = 0$ to $z = 10^{-8}$
also helped to stabilize the results. If seven or eight decimal places of
accuracy were not obtained initially, one of these two adjustments usually
made it possible to do so such that the new result agreed with the initial

result to the same level of accuracy. These adjustments can be viewed as evaluating $F_{1y}{}^A$ at the sample point $(\xi,\eta,0)$ by a limiting procedure. The shifts given above are the smallest shifts that yielded the desired stabilization.

The calculation of the derivatives in Eq. (2.31d) required some care in order to obtain accurate results. Numerical differentiation is inherently unstable as the sample spacing decreases because subtraction of nearly equal function values results in a loss of significance. During the course of performing these calculations, we discovered that a sample spacing of 0.01 was large enough to prevent the occurrence of such numerical instabilities and still yield a sufficiently accurate result for the derivatives. Nevertheless, $\partial^2/\partial\xi^2$ and $\partial^2/\partial\eta^2$ were computed by a five-point finite difference scheme (Abramowitz and Stegun, 1972, 25.3.31) using both 0.01 and 0.02 spacings. We averaged the values from these two calculations. If the two values were different by more than 0.001, a note was made in the data file. As a check on the loss of significance, the number of significant digits was also noted. As an example, the calculation of $\nabla_{\xi\eta}{}^2 F_B{}^A$ exceeded the error tolerance 25 out of 400 times and at least four significant digits were obtained.

APPENDIX C

CALCULATIONS FOR GAUSSIAN BEAM EXPANSION

The Gaussian beam expansion described in Chapter 3 can involve high order Laguerre-Gaussian (LG) modes. Accurate computation of these functions and the associated expansion coefficients for high order is necessary to ensure the success of the method. In this appendix we describe some techniques that achieved the needed accuracy.

Consider the expansion coefficients in Eq. (3.64):

$$D_{pl} = \frac{1}{(|l|+p)!} \left(\frac{|l|}{2} \right) \Gamma \left(p + \frac{|l|}{2} \right) . \tag{1}$$

Evaluating these coefficients directly for large values of p (l fixed) is computationally difficult because of the factor $(|l|+p)!$ in the denominator and the Gamma function in the numerator. These two factors get large very quickly, but the coefficient D_{pl} is dependent on the ratio. In addition, all the coefficients from $p = 0$ to $p = P$ are needed. We addresse these concerns in the recurrence relations described below.

First consider $l =$ even integer. Then the Gamma function can be replaced by a factorial,

$$D_{pl} = \frac{1}{(|l|+p)!} \left(\frac{|l|}{2} \right) (p + \frac{|l|}{2} - 1)! . \tag{2}$$

We can rewrite Eq. (2) as

$$D_{pl} = \left(\frac{|l|}{2} \right) \frac{1}{(p+|l|)(p+|l|-1) \cdots (p+\frac{|l|}{2}+1)(p+\frac{|l|}{2})} . \tag{3}$$

The first coefficient, D_{0l}, is

$$D_{0l} = \left(\frac{1}{2} \right) \frac{1}{(|l|-1)(|l|-2) \cdots (\frac{|l|}{2}+1)(\frac{|l|}{2})} . \tag{4}$$

174

This denominator does not reach an unmanageably high value. The higher order coefficients are found by the following recurrence relation:

$$D_{pl} = D_{(p-1)l} \frac{(p-1+|l|/2)}{(p+|l|)} \quad . \tag{5}$$

Equations (4) and (5) are implemented easily on a computer.

For l = odd integer a similar process yields the relations

$$D_{0l} = \left(\frac{n^{1/2}}{2} \right) \frac{2^{(1-|l|)/2}}{(|l|-1)!!} \quad , \tag{6}$$

$$D_{pl} = D_{(p-1)l} \frac{(|l|+2p-2)/2}{(p+|l|)} \quad , \tag{7}$$

where

$$s!! = \begin{cases} (1)(3)(5) \cdot \cdot \cdot (s-2)(s) & s = 1,3,5,\cdot\cdot\cdot \\ (2)(4)(6) \cdot \cdot \cdot (s-2)(s) & s = 2,4,6,\cdot\cdot\cdot \end{cases} \tag{8}$$

We computed the LG modes by means of the standard recursion relation for the associated Laguerre polynomials (Gradshteyn and Ryzhik, 1980, 8.971.6, 8.973.1 and 8.973.2):

$$L_0^l(\xi) = 1 \quad , \tag{9a}$$

$$L_1^l(\xi) = 1 + l - \xi \quad , \tag{9b}$$

$$(p+1)L_{p+1}^l(\xi) = (2p+l+1-\xi)L_p^l(\xi) + (p+l)L_{p-1}^l(\xi) \quad p = 1,2,3,\cdot\cdot\cdot \tag{9c}$$

The complete LG mode was defined in Eq. (3.20); the part dependent on the normalized radial coordinate is written here in the form

$$\Psi_{pl}(\xi) = \chi^{|l|} L_p^{|l|}(\chi^2) exp(-\chi^2/2) \quad . \tag{10}$$

Notice that in addition to the associated Laguerre polynomial, there are two other factors. One of them is large as $\chi \to \infty$ and the other is small as $\chi \to \infty$. The Gaussian factor dominates for sufficiently large values of χ. Nevertheless, straightforward computation of these factors separately

before they are multiplied together is ill-advised. The algorithm used to compute the LG modes involved multiplying a part of the factors $\chi^{|l|}$ and $exp(-\chi^2/2)$ at every step of the recursive evaluation of $L_p^{|l|}(\chi^2)$.

To understand this, we write these two factors in the following form:

$$\chi^{|l|} \, exp(-\chi^2/2) = exp[-\chi^2/2 + |l| \, ln(\chi)] \quad ,$$

$$\chi^{|l|} \, exp(-\chi^2/2) = exp\left\{(P+1)\frac{[-\chi^2/2 + |l| \, ln(\chi)]}{(P+1)}\right\} \quad ,$$

$$\chi^{|l|} \, exp(-\chi^2/2) = \left(exp\left\{\frac{[-\chi^2/2 + |l| \, ln(\chi)]}{(P+1)}\right\}\right)^{P+1} \quad ,$$

$$\chi^{|l|} \, exp(-\chi^2/2) = \mathcal{A}^{P+1} \quad , \tag{11}$$

where

$$\mathcal{A} = exp\left\{\frac{[-\chi^2/2 + |l| \, ln(\chi)]}{(P+1)}\right\} \quad . \tag{12}$$

In Fig. C.1 we outline the algorithm. At each step the next highest LG mode is computed using Eq. (9). (In Eq. (9) L_p^l is replaced with the current value of ψ_{pl}.) Then that mode and all the preceding modes are multiplied by the factor $\mathcal{A}$. There are $P+1$ step overall (modes $p = 0, 1, 2,$ $\cdots, P$). Thus, the factor $\mathcal{A}$ is multiplied into each mode $P+1$ times and the computational limit on χ for calculating the modes is extended.

We checked the numerical stability and accuracy in a number of ways. To supplement test calculations comparing computed results to tabulated results (agreement was obtained to the accuracy of the tabulated results), a finite sum test was used. The summation relation for associated Laguerre functions (Gradshteyn and Ryzhik, 1980, 8.974.3),
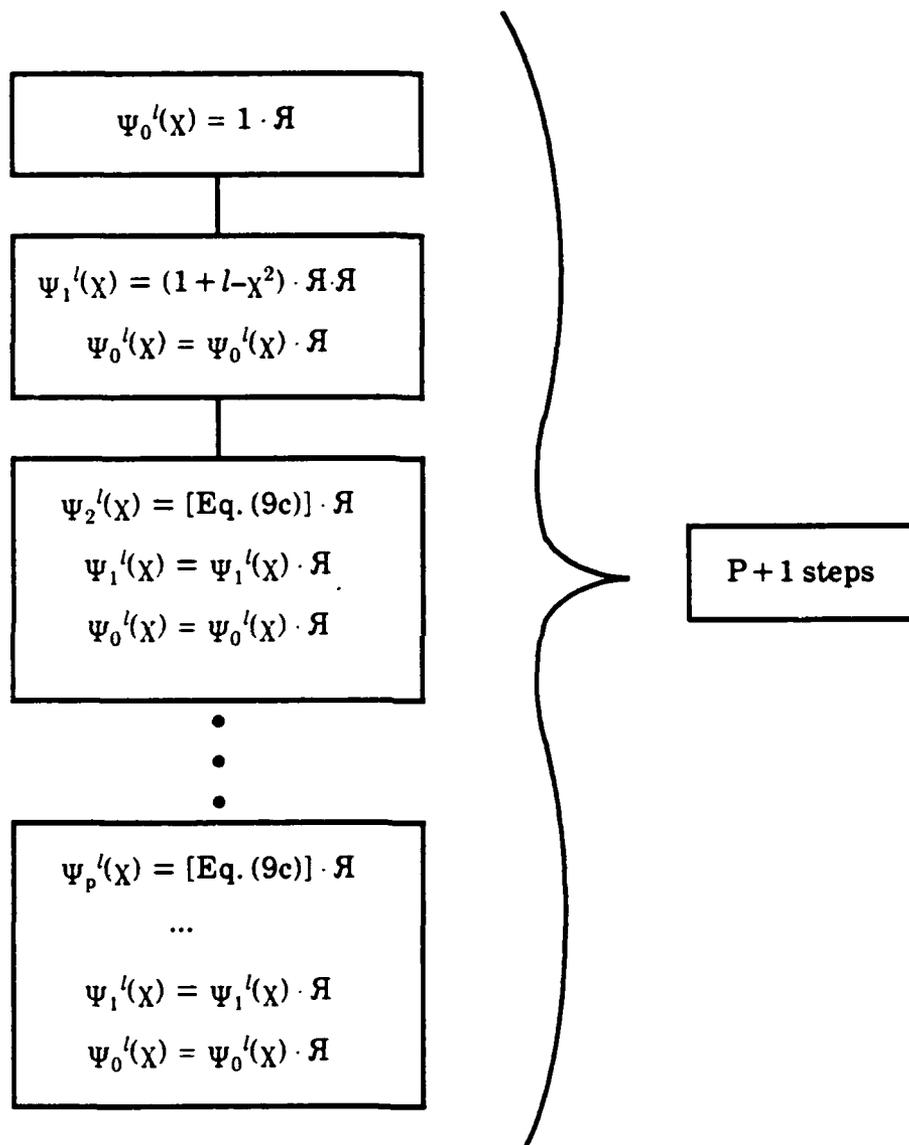
Fig. C.1. Algorithm for computation of LG modes.

$$\sum_{p=0}^{P} L_p^l(\xi) = L_p^{l+1}(\xi) \quad , \tag{13}$$

was generalized for LG modes, i.e.,

$$\sum_{p=0}^{P} \chi^{|l|} L_p^{|l|}(\chi^2) \, exp(-\chi^2/2) = \chi^{|l|} L_p^{|l|+1}(\chi^2) \, exp(-\chi^2/2) \quad . \tag{14}$$

Equation (14) was satisfied to seven decimal places for large values of p, l, and χ. In addition, since computation of the LG modes in double precision agreed with their computation in single precision, we deduced that the recursion algorithm is stable over the parameter ranges of interest.

APPENDIX D

QUANTITATIVE DETERMINATION OF OPTICAL INTENSITY

FROM PHOTOGRAPHIC FILM

In this appendix we describe in detail some of the equipment used in making the experimental measurements of Chapter 4. Figure 4.20 contained a block diagram of the experimental steps followed. Also, we explain how quantitative measurements of the optical intensity were deduced from measurements of the film grain density.

We recorded the diffraction patterns on KODAK T-MAX 100 film. T-MAX 100 "is a continuous-tone panchromatic black-and-white negative film for general outdoor and indoor photography; it is especially useful for detailed subjects when you need maximum image quality." (Eastman Kodak Co., pamphlet F-25) The film offered medium speed (EI 100), fine grain, improved reciprocity, and an expanded exposure latitude. This last characteristic was particularly advantageous in recording a wide range of intensity data on a single frame.

In order to achieve a film grain density ≥ 3 and to compress some of the broad exposure latitude into this density range, we developed the film for a 40% increase in contrast above normal. With KODAK HC-110 developer (Dil B), an increase in development time of 80% at 20°C was necessary. As a result the density range of the film was well matched to the density sensitivity of the P-1700 scanner. The processing details are listed in Table D.1.

Table D.1. KODAK T-MAX 100 Process Times
at 20°C.

Process Step	Time
KODAK HC-110 (Dil B) (agitate every 30 sec.)	12.5 min.
KODAK Indicator Stop Bath	30 sec.
KODAK Rapid Fixer (agitate every 30 sec.)	3 min.
Wash in running water	5-10 min.
KODAK PHOTO-FLO	30 sec.
Dry	---

We scanned and digitized the recorded diffraction patterns using image processing instrumentation by Optronics International, Inc. (model Photomation P-1700 scanner/plotter). Only the scanning capabilities of the device were used. The P-1700 is described in the *Operation and Maintenance Manual* (Optronics, 1984). The following description is taken from that source (pp. 1-1 and 1-2):

> [The Photomation P-1700] is a high-speed digital scanner that utilizes electronic control and mechanical operation to optically scan an image mounted on a rotating drum. Photometric data from the scanned image is rapidly converted to digital form for computer processing and storage. A scan medium (film transparency or opaque material) is clamped onto the drum so that the scan medium conforms to the drum's machined cylindrical surface. A Koehler illumination system ensures uniform illumination and the focusing of turret-mounted apertures on the medium surface. Light transmitted through the film (or reflected from the opaque material) is measured by a photodetector.

> The effective spot size for the illuminating light is selectable from the following sizes: 15, 30, 60, 120, 240, and 480 microns square. The receiving (imaging) optical apertures are selectable at 12.5, 25, 50, 100, 200 or 400 microns square.

> The illuminating and imaging optics are mounted on opposite arms of a C-shaped carriage, through which the drum rotates. Optical

density of the film along the circumference of the drum (y direction) is measured every 12.5, 25, 50, 100, 200 or 400 microns. After each rotation, the optical carriage is stepped in the axial direction by the selected raster size, and the process is repeated until the total area of interest is scanned. Therefore, any rectangular area of a film, the limits of which are selectable by the operator, can be scanned with a line separation of 12.5, 25, 50, 100, 200 or 400 microns – depending upon the aperture selected. An opening in the drum provides for the measurement of light transmission through air. The densitometer photodetctor (sic) system is recalibrated by this measurement once each drum rotation.

The detector current, resulting from the light transmitted through or reflected from the scan medium, is amplified logarithmically and digitized. A word length of eight bits is used, allowing for a possible $2^8 = 256$ unique words. The words represent the densities of the picture elements (pixels) scanned relative to the density range selected; thus, a number of 255 would be equivalent to an optical density of 2 for a sensitivity range of 0 to 2D, or to an optical density of 3 for a sensitivity range of 0 to 3D.

Each of the fourteen frames (1/2000, 1/1000, ···, 2, 4 seconds) of

the diffraction pattern was scanned. We selected raster size 50 μm; thus, a

512×512 scan covered 25.6 mm $\times$ 25.6 mm. This scan area was large

enough to include the first three diffracted orders and small enough to

yield a manageable number of digitized data points. The 512×512 format

was appropriate for most of the subsequent processing.

Because we developed the film to high contrast, a sensitivity range of

0 to 3D was chosen. Thus, pixel value 255 corresponds to an optical density

of 3. As a result a single frame mapped intensity in the diffraction pattern

over three orders of magnitude into pixel values 0–255. Varying the

exposure time from 1/2000 second to 4 seconds added almost four more

orders of magnitude to the dynamic range of the intensities recorded.

The resources of the International Imaging Systems, Inc. image

processing system (IIS) made many of the data manipulations,

registrations, and displays easy. The image processing system consisted of

the Model 75 Image Processor and the System 600 Programming Environment. "The Model 75 Image Processor is an advanced, general purpose image array processor with unique computation features which enhance the performance of the digital image processing system. In addition to the color display of imagery, the Model 75 provides the architecture to implement a wide variety of complex image processing algorithms." (International Imaging Systems, Inc., Model 75 Image Processor Product Description, 1983). The Model 75 was hosted by a MASSCOMP 5700 computer (68020 processor based, 4 Mb, Unix AT&T SYS V.2 operating system). The System 600 Programming Environment controlled the Model 75 and included a large number of image processing functions. Image display, statistical computation, registration, and piecewise linear mapping of pixel values were among the most used procedures.

In Fig. 4.21 we showed a simplified block diagram of the experimental steps. The net effect of the film–scanner step was to apply a D-log-e curve to the diffraction pattern intensity (D is optical density and e is exposure). The output of the scanner was a pixel value (PV) from 0–255, however, not optical density (although there is a direct relationship). More appropriately, then, the film–scanner step applied a PV-log-e curve to the diffraction pattern intensity. The output from the scanner was displayed and analyzed by the image processing software and hardware.

Thus, it was necessary to determine the PV-log-e curve for the film used. The following procedure achieved this. The camera with a 50 mm, f/1.2 lens was mounted above a copy stand light table upon which were

placed neutral density filters of various values (0.0, 0.3, 0.6, 1.0, 1.3, 1.6, and 2.0). We recorded photographs of the test scene at every shutter speed (1/2000, 1/1000, ⋯, 2, 4, 8 seconds); the camera metered correct exposure at 1/125 second. (Note: to eliminate variations in development from affecting results, we used the first half of a roll of film to take the PV-log-e test scene photographs and the second half of the roll to take the experimental diffraction pattern photographs.)

After development, we scanned every test scene on the P-1700 and analyzed the data using the IIS software. The analysis consisted of performing a statistical average and standard deviation of the pixel values over the regions of interest, i.e., the images of the neutral density filters. These values were recorded and plotted versus the logarithm of exposure. Arbitrarily, a neutral density of 0.0 was assigned unit intensity ($EE^* = 1.0$), and other neutral densities corresponded to lower intensities [$EE^*(0.3) = -0.3$, $EE^*(0.6) = -0.6$, and so on]. The exposure time of the frame was used to compute e.

The data are shown in Fig. D.1. The range of exposure times and neutral density values was sufficient to cover the film from fog (PV = 13) to saturation (PV = 255). In addition, data from adjacent frames overlapped greatly, thereby showing internal consistency among the data.

We decided to fit a piecewise linear curve to approximate the PV-log-e curve of the film and the data of Fig. D.1. The curve would have zero slope to the toe of the data, a positive slope along the first linear part of the data, a positive slope along the second linear part of the data, and zero slope from the shoulder of the data. This curve, determined by a specialized least-

squares procedure, is

$$
PV = \begin{cases}
13 & \log e \le -3.166 \\
72.525\ \log e + 242.650 & -3.166 \le \log e \le -1.602 \\
114.327\ \log e + 309.616 & -1.602 \le \log e \le -0.478 \\
255 & -0.478 \le \log e
\end{cases} \tag{1}
$$

The mean square error of Eq. (1) from the actual data is ± 5.070; hence, the
uncertainty in determining the pixel value (or the optical density) given
the logarithm of exposure is about 2%.

Using Eq. (1) as a model for the PV-log-e curve of the film, we made
an analysis of the experimental photographs of §4.2.3. The analysis
involved using the PV-log-e curve to deduce intensity values from the
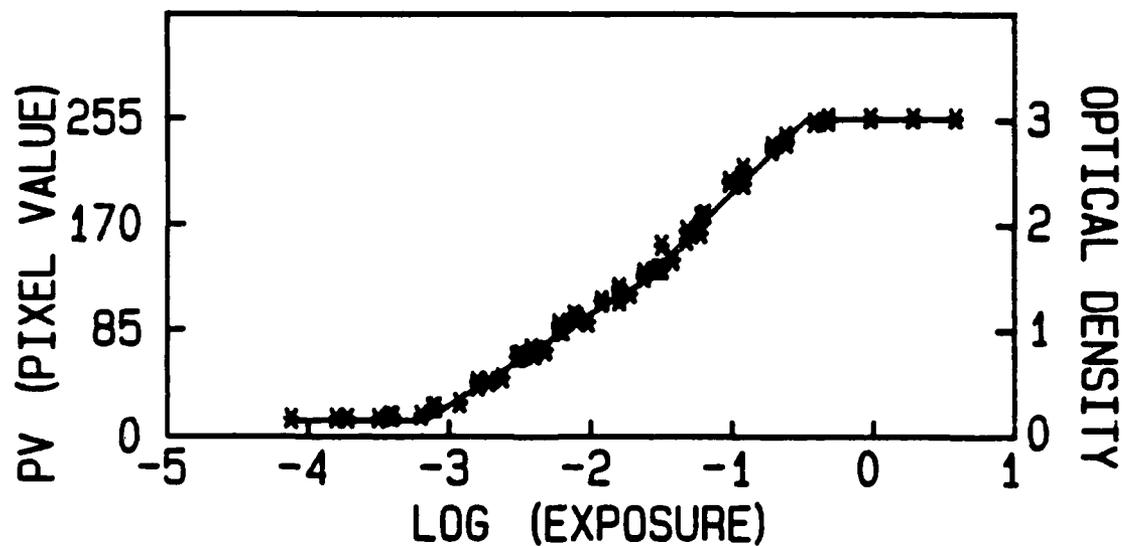
Fig. D.1. PV-log-e (pixel value vs logarithm of exposure) curve for T-MAX
100 film (high contrast). The experimental data points are indicated by
stars ($*$) and span the pixel range from fog (PV = 13) to saturation (PV =
255). A piecewise linear fit to the data is shown as a solid line (———). The
corresponding optical density is shown on the right-hand vertical axis.

experimental photographs. Several photographs were sampled along a line and compared with the theoretical computations along that same line.

To deduce the normalized intensity, $E_0E_0{}^*$, from Eq. (1), we find the inverse function. Expressing this inverse functions first in terms of $\log e$ gives

$$
\log e = \begin{cases}
? = -3.166 & 0 \le PV \le 13 \\
(PV - 242.650)/\ 72.525 & 14 \le PV \le 126 \\
(PV - 309.616)/114.327 & 127 \le PV \le 254 \\
? = -0.478 & 255 \le PV
\end{cases} \tag{2}
$$

Note that the zero slope sections of the PV-log-e curve mean that $\log e$ is indeterminate from PV there. We assigned the breakpoint values of –3.166 (corresponding to $PV \le 13$) and –0.478 (corresponding to $PV = 255$) for these regions.

We wanted to deduce the logarithm of the intensity normalized in the same way as the computational results, i.e., normalized to the central peak value. Calling this normalizing value I_0, the logarithm of the normalized intensity, $E_0E_0{}^*$, is

$$
\log E_0E_0^{\bullet} = \log \frac{EE^{\bullet}}{I_0} = \log e - \log t - \log I_0 \tag{3}
$$

The value for $\log e$ can be determined by Eq. (2), and $\log t$ is a number dependent on the shutter speed. We determined I_0 in the following way.

Examination of the 1/500 second photograph revealed that the PV of the first order peak was 129 and the PV of the second order peak was 73. The computational result predicted that these peaks would have normalized intensities of $10^{-2.251}$ and $10^{-2.910}$, respectively. The value of $\log I_0$ that approximately produced these normalized intensities given the observed PV in the experimental photographs is

$$log\ I_0 = 3.320 \qquad (4)$$

The computational model for the film–scanner step is complete. The pixel values measured from the scanning step were converted to normalized intensity values by the following two steps: 1) determine $log\ e$ from Eq. (2) and 2) compute $log\ E_0 E_0^*$ using Eq. (3), the value for exposure time, and the value from Eq. (4).